

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:27:18 ; Search time 169,442 Seconds
(without alignments)
8602.339 Million cell updates/sec

Title: US-09-513-151-3_COPY_1121_1210

Perfect score: 90
Sequence: 1 TGTGACCTCTGTGATCGAAT.....TGACCACTGAAGAAAGA 90

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estda:*
2: em_esthm:*
3: em_estnu:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	90	100.0	562	12	BG495857 602540248
2	90	100.0	711	12	BG506209 601860153
3	90	100.0	724	14	BM721352 UI-E-E01-
4	90	100.0	772	12	BG034660 602300022
5	90	100.0	869	12	BG612651 602640078
6	90	100.0	1014	14	BQ221144 AGENCOURT

7	90	100.0	1088	14	BM800217	AGENCOURT
8	88.4	98.2	674	10	AV685706	AV685706
9	86	93.6	579	14	W27852	4143 Human
10	84.4	93.8	987	9	AL530574	AL530574
11	78.8	87.6	485	10	BE477083	BE477083
12	78.6	87.3	922	9	AL530575	AL530575
13	78	86.7	724	12	BF207826	BF207826
14	74	82.2	468	10	BE332036	BE332036
15	74	82.2	557	12	BE851413	BE851413
16	74	82.2	762	13	BI685252	603310265
17	74	82.2	766	13	BI107517	BI107517
18	74	82.2	780	14	BQ571052	UI-M-FBO-
19	74	82.2	2061	11	AK003556	AK003556
20	73	81.1	785	12	BF133971	BF133971
21	72.4	80.4	1951	11	AK010176	AK010176
22	61.8	68.7	242	9	AV218623	AV218623
23	51.4	57.1	336	10	AV695603	AV695603
24	48.2	53.6	836	12	BF185342	BF185342
25	46	51.1	438	9	AL698051	AL698051
26	43.2	48.0	632	13	BJ026001	BJ026001
27	42.4	47.1	718	13	BJ016230	BJ016230
28	38.4	42.7	178	14	BM798211	BM798211
29	36.9	44.2	442	10	BE531754	BE531754
30	33.2	35.3	479	17	A2054089	A2054089
31	31.8	34.4	473	17	GC0000548	GC0000548
32	31	34.2	513	13	BI808143	BI808143
33	30.8	34.2	700	12	BG827372	BG827372
34	30.2	33.6	1101	17	CNS00K50	CNS00K50
35	29.8	33.1	666	10	BE261674	BE261674
36	29.4	32.7	1052	17	CNS010HY	CNS010HY
37	29	32.2	354	13	BI062622	BI062622
38	28.4	31.6	758	13	BI686281	BI686281
39	28.2	31.3	396	10	AM024086	AM024086
40	28.2	31.3	439	14	BU024120	BU024120
41	28.2	31.3	512	10	AM509578	AM509578
42	28.2	31.3	525	9	AU241057	AU241057
43	28.2	31.3	590	13	BJ007993	BJ007993
44	28.2	31.3				
45	28.2	31.3				

ALIGNMENTS

RESULT 1
BG495857
LOCUS 562 bp mRNA linear EST 27-MAR-2001
DEFINITION 602540248F1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4671566 5',
mRNA sequence.
ACCESSION BG495857
VERSION BG495857.1 GI:13457373
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 562)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNC4186 row: 0 column: 11
High quality sequence stop: 562.
Location/Qualifiers
1. 562

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4671586"
/clone_lib="NIH_MGC_59"
/tissue_type="mucoepidermoid carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgcttggcc); Site_2: SfiI (ggccattatggcc);
Double stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATTATAGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCCAGCGCCGACATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT      167 a      122 c      142 g      131 t
ORIGIN
Query Match      100.0%; Score 90; DB 12; Length 562;
Best Local Similarity 100.0%; Pred. No. 8.9e-20;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TGTGACCTCTGTGATCGAATCATCTTGGGATCGCGAATGCGCAGCCACATAAATCC 60
|||||
Db 205 TGTGACCTCTGTGATCGAATCATCTTGGGATCGCGAATGCGCAGCCACATAAATCC 264
OY 61 AATCCCACTTGAAACCACTGAAGAAAGA 90
|||||
Db 265 AATCCCACTTGAAACCACTGAAGAAAGA 294

RESULT 2
LOCUS      BG506209      711 bp      mRNA      linear      EST 27-MAR-2001
DEFINITION      601860153P1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4072428 5',
mRNA sequence.
ACCESSION      BG506209
VERSION      BG506209.1 GI:13467726
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1 (bases 1 to 711)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      Unpublished (1999)
JOURNAL
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgabbs@email.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM919 row: j column: 13
High quality sequence stop: 643.
Location/Qualifiers
1..711
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4072428"
/clone_lib="NIH_MGC_61"
/tissue_type="embryonal carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgcttggcc); Site_2: SfiI (ggccattatggcc);
Double stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
```

```
adaptor sequence: 5'-CACGGCCATTATAGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCCAGCGCCGACATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT      224 a      145 c      165 g      177 t
ORIGIN
Query Match      100.0%; Score 90; DB 12; Length 711;
Best Local Similarity 100.0%; Pred. No. 9.4e-20;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TGTGACCTCTGTGATCGAATCATCTTGGGATCGCGAATGCGCAGCCACATAAATCC 60
|||||
Db 74 TGTGACCTCTGTGATCGAATCATCTTGGGATCGCGAATGCGCAGCCACATAAATCC 133
OY 61 AATCCCACTTGAAACCACTGAAGAAAGA 90
|||||
Db 134 AATCCCACTTGAAACCACTGAAGAAAGA 163

RESULT 3
LOCUS      BM721352      724 bp      mRNA      linear      EST 01-MAR-2002
DEFINITION      UI-E-E01-1b-b-20-0-UI-r1 UI-E-E01 Homo sapiens cDNA clone
UI-E-E01-1b-b-20-0-UI 5', mRNA sequence.
ACCESSION      BM721352
VERSION      BM721352.1 GI:19041207
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1 (bases 1 to 724)
AUTHORS      Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
COMMENT      Contact: Soares, MB
9704447
University of Iowa
Program for Rat Gene Discovery and Mapping
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msquares@blue.weeg.iowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1..724
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-E01-1b-b-20-0-UI"
/clone_lib="UI-E-E01"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pRT3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-E01 is a normalized cDNA library containing the
following tissue(s): fetal eye. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
```


adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCGCATACC. This library was created for the program Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 208 a 132 c 175 g 208 t 1 others

ORIGIN

Query Match 100.0%; Score 90; DB 14; Length 724;
Best Local Similarity 100.0%; Pred. No. 9.4e-20;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGGCACATATAATCC 60
|||||
Db 5 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGGCACATATAATCC 64
|||||

OY 61 AAATCCCACTTGAACCACTGAAGAAAAGA 90
|||||
Db 65 AAATCCCACTTGAACCACTGAAGAAAAGA 94
|||||

RESULT 4
BG034660 772 bp mRNA linear EST 24-JAN-2001
LOCUS 602300022F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4394260 5',
DEFINITION mRNA sequence.
ACCESSION BG034660
VERSION BG034660.1 GI:12428227
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 772)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: LHAM0089 row: 1 column: 05
High quality sequence stop: 692.
Location/Qualifiers
1. 772

FEATURES
source

BASE COUNT 224 a 155 c 178 g 214 t 1 others

ORIGIN

Query Match 100.0%; Score 90; DB 12; Length 772;
Best Local Similarity 100.0%; Pred. No. 9.5e-20;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGGCACATATAATCC 60
|||||
Db 181 AAATCCCACTTGAACCACTGAAGAAAAGA 210
|||||

Db 121 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGGCACATATAATCC 180
OY 61 AAATCCCACTTGAACCACTGAAGAAAAGA 90
|||||
Db 181 AAATCCCACTTGAACCACTGAAGAAAAGA 210
|||||

RESULT 5
BG612651 869 bp mRNA linear EST 18-APR-2001
LOCUS 602640078F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4771338 5',
DEFINITION mRNA sequence.
ACCESSION BG612651
VERSION BG612651.1 GI:13664022
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 869)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: LHCMI639 row: k column: 19
High quality sequence stop: 684.
Location/Qualifiers
1. 869

FEATURES
source

BASE COUNT 259 a 184 c 206 g 220 t

ORIGIN

Query Match 100.0%; Score 90; DB 12; Length 869;
Best Local Similarity 100.0%; Pred. No. 9.8e-20;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGGCACATATAATCC 60
|||||
Db 74 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGGCACATATAATCC 133
|||||

OY 61 AAATCCCACTTGAACCACTGAAGAAAAGA 90
|||||
Db 134 AAATCCCACTTGAACCACTGAAGAAAAGA 163
|||||

RESULT 6
BQ221144 1014 bp mRNA linear EST 02-MAY-2002
LOCUS BQ221144

DEFINITION	AGENCOURT_7590416 NIH_MGC_72 Homo sapiens CDNA clone IMAGE:6065448									
ACCESSION	5', mRNA sequence.									
VERSION	BQ221144									
KEYWORDS	BQ221144.1 GI:20402544									
SOURCE	EST.									
ORGANISM	human.									
	Homo sapiens									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.									
REFERENCE	1 (bases 1 to 1014)									
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .									
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)									
JOURNAL	Unpublished (1999)									
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC/DCTO/DRP CDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Agencourt Bioscience Corporation clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LRAM1341 row: m column: 01 High quality sequence stop: 387. Location/Qualifiers									
FEATURES	1..1014									
SOURCE	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:6065448" /clone_1lb="NIH.MGC.72" /tissue_type="melanotic melanoma" /lab_host="DH10B (phage-resistant)" /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."									
BASE COUNT	289 a 283 c 231 g 211 t									
ORIGIN										
Query Match	100.0%; Score 90; DB 14; Length 1014;									
Best Local Similarity	100.0%; Pred. No. 1e-19;									
Matches	90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Oy	1 TGTGACCTTGATGATCATCATGTTGGGATGCGAATGGCAGCCACATAAATCC 60 									
Db	295 TGTGACCTTGATGATCATCATGTTGGGATGCGAATGGCAGCCACATAAATCC 354 									
Oy	61 AATCCCACTTGACCACTGAAGAAGA 90 									
Db	355 AATCCCACTTGACCACTGAAGAAGA 384 									
RESULT 7										
LOCUS	BM800217 1088 bp mRNA linear EST 05-MAR-2002									
DEFINITION	AGENCOURT_6416157 NIH_MGC_71 Homo sapiens CDNA clone IMAGE:5531248									
ACCESSION	5', mRNA sequence.									
VERSION	BM800217									
KEYWORDS	BM800217.1 GI:19117040									
SOURCE	EST.									
ORGANISM	human.									
	Homo sapiens									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.									
REFERENCE	1 (bases 1 to 1088)									
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .									
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)									
JOURNAL	Unpublished (1999)									
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)									

FEATURES									
source									
1. 1088									
/organism="Homo sapiens"									
/db_xref="taxon:9606"									
/clone="IMAGE:5531248"									
/clone_1id="NH_MGC_71"									
/tissue_type="telomysarcoma"									
/lab_host="DH10B (phage-resistant)"									
/note="Organ: uterus; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.									
Average insert size 2.1 kb.									
BASE COUNT 318 a 218 c 255 g 294 t 3 others									
ORIGIN									
Query Match 100.0%; Score 90; DB 14; Length 1088;									
Best Local Similarity 100.0%; Pred. NO. 1e-19;									
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
OY 1 TGTGACCTCTGATGCATCATCATTCGGATGCGGATGCGAGCGCACATAAATCC 60									
DB 283 TGTGACCTCTGATGCATCATCATTCGGATGCGGATGCGAGCGCACATAAATCC 342									
OY 61 AATCCCACTTGACCACTGAAGAAGA 90									
DB 343 AATCCCACTTGACCACTGAAGAAGA 372									
RESULT 8									
AV685706 674 bp mRNA linear EST 16-JAN-2002									
LOCUS AV685706 GKC Homo sapiens cDNA clone GKCBYG04 5', mRNA sequence.									
DEFINITION AV685706									
ACCESSION AV685706.1 GI:10287569									
VERSION EST.									
KEYWORDS human.									
SOURCE									
ORGANISM Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.									
REFERENCE 1 (bases 1 to 674)									
XU.X., HUANG.J., XU.Z., QIAN.B., ZHU.Z., YAN.Q., CAI.T., ZHANG.X.,									
XIAO.H., QU.D., LIU.F., HUANG.Q., CHENG.Z., LI.N., DU.J., HU.W.,									
SHEN.K., LU.G., FU.G., ZHONG.M., XU.S., GU.W., HUANG.W., ZHAO.X.,									
HU.G., GU.J., CHEN.Z. and HAN.Z.									
Insight into hepatocellular carcinogenesis at transcriptome level									
by comparing gene expression profiles of hepatocellular carcinoma									
with those of corresponding noncancerous liver									
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)									
COMMENT									
JOURNAL MEDLINE									
Contact: Zeguang Han									
Chinese National Human Genome Center at Shanghai									
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai									
201203, P. R. China									
Tel: 86-21-50801919(ex. 45)									
Fax: 86-21-50801922									
Email: hanzgchc.sh.cn									
This clone is available at CHGC in Shanghai.									
FEATURES									
source									
1. 674									
/organism="Homo sapiens"									
/db_xref="taxon:9606"									
/clone="GKCBYG04"									
/clone_1id="GKC"									
/tissue_type="hepatocellular carcinoma"									
/dev_stage="Adult"									
/lab_host="SOLR"									
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:									
XhoI"									

```

BASE COUNT      206 a      141 c      163 g      157 t      7 others
ORIGIN
Query Match      98.2%; Score 88.4; DB 10; Length 674;
Best Local Similarity 98.9%; Pred. No. 3.2e-19;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATGATCATCATTTGGGATCGCGAATGGCGAGCCACATATAATCC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 323 TGTGACCCCTGTGATGATCATCATTTGGGATCGCGAATGGCGAGCCACATATAATCC 382
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 61 AAATCCCACTTGAACCACTGAAGAAAAGA 90
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 383 AAATCCCACTTGAACCACTGAAGAAAAGA 412
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
W27852      579 bp      mRNA      linear      EST 08-MAY-1996
LOCUS       W27852/C
DEFINITION  41d3 Human retina cDNA randomly primed sublibrary Homo sapiens cDNA
            mRNA sequence.
ACCESSION   W27852
VERSION     W27852.1 GI:1308007
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 579)
AUTHORS     Macke, J., Smallwood, P. and Nathans, J.
TITLE       Adult Human Retina cDNA
JOURNAL     Unpublished (1996)
COMMENT     Contact: Dr. Jeremy Nathans
            Dr. Jeremy Nathans, Dept. of Molecular Biology and Genetics
            Johns Hopkins School of Medicine
            725 North Wolfe Street, Baltimore, MD 21205
            Tel: 410 955 4678
            Fax: 410 614 0827
            Email: jeremy.nathans@jhu.edu
            Clones from this library are NOT available.
PCR PRIMERS
FORWARD: CTTTGGAGCAAGTTGACGCTGTTAACT
BACKWARD: GAGTGGCTTACGATGATTTCTCCAGCGTAA
Seq primer: GGGTAAAGCAAGAAAGAT.

FEATURES
    source
        1..579
            location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="Human retina cDNA randomly primed sublibrary"
                /sex="mixed (males and females)"
                /tissue_type="retina"
                /dev_stage="adult"
                /lab_host="E. coli strain K802"
                /note="Organ: eye; Vector: lambda gt10; Site_1: EcoRI;
                Site_2: EcoRI; The library used for sequencing was a
                sublibrary derived from a human retina cDNA library.
                Inserts from retina cDNA library DNA were isolated.
                Randomly primed, PCR amplified, size-selected, and cloned
                into lambda gt10. Individual plaques were arrayed and used
                as templates for PCR amplification, and these PCR
                products were used for sequencing."

BASE COUNT      97 a      110 c      91 g      136 t      145 others
ORIGIN
Query Match      95.6%; Score 86; DB 14; Length 579;
Best Local Similarity 96.6%; Pred. No. 2e-18;
Matches 86; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

DB 258 NAATCCCACTTGAACCACTGAGAAAAG 230
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
AL530574/C      987 bp      mRNA      linear      EST 13-FEB-2001
LOCUS       AL530574/C
DEFINITION  AL530574 LTI_NFL001_NBC4 Homo sapiens cDNA clone CSDD008YN11 3
            prime, mRNA sequence.
ACCESSION   AL530574
VERSION     AL530574.1 GI:12794067
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 987)
AUTHORS     Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 Evry cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
    source
        1..987
            location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="CSDD008YN11"
                /clone_lib="LTI_NFL001_NBC4"
                /sex="male"
                /tissue_type="neuroblastoma cells"
                /lab_host="DH10B"
                /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                was primed with a NotI-oligo(dT) primer. Five prime end
                enriched, double-stranded cDNA was digested with Not I and
                cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                vector. Library was normalized. Library was constructed
                by Life Technologies. Contact : Feng Liang Life
                Technologies, a division of Invitrogen 9800 Medical Center
                Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                8371 Email : fliang@life.com"
                http://fulllength.invitrogen.com"

BASE COUNT      281 a      230 c      180 g      284 t      12 others
ORIGIN
Query Match      93.8%; Score 84.4; DB 9; Length 987;
Best Local Similarity 94.4%; Pred. No. 7.6e-18;
Matches 85; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATGATCATCATTTGGGATCGCGAATGGCGAGCCACATATAATCC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 885 TGTGACCTCTGTGATGATCATCATTTGGGATCGCGAATGGCGAGCCACATATAATCC 826
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 61 AAATCCCACTTGAACCACTGAAGAAAAGA 90
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 825 AAATCCCACTTGAACCACTGAAGAAAAGR 796
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
BE477083      485 bp      mRNA      linear      EST 28-AUG-2000
LOCUS       BE477083
DEFINITION  160729 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BE477083
VERSION     BE477083.1 GI:9596688
KEYWORDS    EST.
SOURCE      cow.
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bos.
            Sonstegard, T.S., Capuco, A.V., Van Tassel, C.P., Ashwell, M.S. and
            1 (bases 1 to 485)

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 19:19:26 ; Search time 5959 Seconds

(Without alignments)
9967.919 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041
Sequence: 1 CTCGCATTAAGATGCGCTCCG.....TTTACACAGAAAAA 2041Scoring table: OLIGO_NUC
Gapex 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

```
GenEmbl: *
1: gb_ba: *
2: gb_hlg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_hlg_hum: *
31: em_hlg_inv: *
32: em_hlg_oth: *
33: em_hlg_mus: *
34: em_hlg_pln: *
35: em_hlg_rtd: *
36: em_hlg_mam: *
37: em_hlg_vrt: *
38: em_sy: *
39: em_hlggo_hum: *
40: em_hlggo_mus: *
41: em_hlggo_other: *
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB	ID	Description
1	2041	100.0	2041	6	A98668
2	1258	61.6	1308	6	AY052768 Homo sapi
3	896	43.9	2130	6	AY405704 Sequence
4	825	40.4	2216	9	AF074918 Homo sapi
5	765	37.5	1844	9	BC010741 Homo sapi
6	751	36.8	1749	9	AK000068 Homo sapi
7	751	36.8	1749	9	AK074222 Homo sapi
8	593	29.1	166496	9	HS118121 Human DNA
9	593	29.1	200484	2	AC025002 Homo sapi
10	473	23.2	199863	2	AL354888 Homo sapi
11	473	23.2	199863	2	AL354888 Homo sapi
12	241	11.8	355	6	AX336538 Sequence
13	150	7.3	252	11	G24438
14	143	5.9	199863	2	AL354888 Homo sapi
15	120	5.9	79785	2	AC025068 Homo sapi
16	48	2.4	2045	10	BC019812 Mus muscu
17	48	2.4	198772	10	AL606906 Mouse DNA
18	34	1.7	77304	2	AC103185
19	34	1.7	201861	2	AC103344 Rattus no
20	32	1.6	201861	2	AC103344 Rattus no
21	26	1.3	66680	2	AC100483 Mus muscu
22	26	1.3	69888	2	AC087576 Homo sapi
23	26	1.3	75021	2	AC040889 Homo sapi
24	26	1.3	129332	2	AL590028 Human DNA
25	26	1.3	145363	2	AC079213 Homo sapi
26	26	1.3	149646	2	AL163536 Homo sapi
27	26	1.3	160708	2	AC068406 Homo sapi
28	26	1.3	161486	2	AL590008 Homo sapi
29	26	1.3	163603	2	AC026755 Homo sapi
30	26	1.3	195169	2	AC017109 Homo sapi
31	26	1.3	197310	9	AC099754 Homo sapi
32	26	1.3	198958	2	AC020688 Homo sapi
33	25	1.2	250529	9	HUAE000658 Homo sapi
34	24	1.2	91767	9	AC007748 Homo sapi
35	24	1.2	99593	9	HS085717 Human DNA
36	24	1.2	157658	2	AC073285 Homo sapi
37	24	1.2	158692	9	AC009490 Homo sapi
38	24	1.2	162548	2	AC092659 Homo sapi
39	24	1.2	163377	2	AC131158 Homo sapi
40	24	1.2	166588	2	AC007897 Homo sapi
41	24	1.2	166848	2	AC021870 Homo sapi
42	24	1.2	171619	9	AP002367 Homo sapi
43	24	1.2	175822	2	AC008082 Homo sapi
44	24	1.2	192338	2	AC068402 Homo sapi
45	24	1.2	193749	2	AC022221 Homo sapi
					AC087286 Homo sapi

ALIGNMENTS

```
RESULT 1
A98668 LOCUS A98668 2041 bp DNA linear PAT 26-JAN-2000
DEFINITION Sequence 3 from Patent WO9910482.
ACCESSION A98668
VERSION A98668.1 GI:6781710
KEYWORDS
ORGANISM
SOURCE
REFERENCE 1 (bases 1 to 2041)
AUTHORS Lemieux, J. and Hekimi, S.
TITLE THE C. ELEGANS GRO-1 GENE
JOURNAL Patent: WO 9910482-A 3 04-MAR-1999;
LEMIEUX JASON (CA); UNIV MCGILL (CA)
```

FEATURES
 source 1.2041 Location/Qualifiers
 /organism="unidentified"
 /db_xref="taxon:3264"
 BASE COUNT 589 a 421 c 502 g 529 t
 ORIGIN

Query Match 100.0%; Score 2041; DB 6; Length 2041;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCCATTAAGATGCGCTGCGCGCTGCAGAGCAGTTCTGTGGCGAGTGGGCTCAG 60
 |||||
 Db 1 CTGCCATTAAGATGCGCTGCGCGCTGCAGAGCAGTTCTGTGGCGAGTGGGCTCAG 60

QY 61 GGGCTGCAAGGAGACCTTACCTCTTGTAGTATCTCGGGCCACGGGCAACCAATC 120
 |||||
 Db 61 GGGCTGCAAGGAGACCTTACCTCTTGTAGTATCTCGGGCCACGGGCAACCAATC 120

QY 121 CAGGCTGGGCTTGCAGCTAGGCGGCGCTCGGCGGTGAGATGTCAGCGCTGACTCAT 180
 |||||
 Db 121 CAGGCTGGGCTTGCAGCTAGGCGGCGCTCGGCGGTGAGATGTCAGCGCTGACTCAT 180

QY 181 GCAGGTCTATGAAGCGCTTAGACATCATCACCAAGGTTTGTGCCAAGAGCAGAGAT 240
 |||||
 Db 181 GCAGGTCTATGAAGCGCTTAGACATCATCACCAAGGTTTGTGCCAAGAGCAGAGAT 240

QY 241 CTGCGGCGACACATGATCAGCTTTGTGATCTCTTGTGACCAATTCACAGTGTGGA 300
 |||||
 Db 241 CTGCGGCGACACATGATCAGCTTTGTGATCTCTTGTGACCAATTCACAGTGTGGA 300

QY 301 CTTTCAAGAAATAGACAACTCTCTGATGTAAGATATATTGCCCCGAGACAAATTCCTAT 360
 |||||
 Db 301 CTTTCAAGAAATAGACAACTCTCTGATGTAAGATATATTGCCCCGAGACAAATTCCTAT 360

QY 361 TGTGTGGGAGGAGCAACATATTACATTGTAATCTGTGCTGTGAAAGTTTGTGCAATAC 420
 |||||
 Db 361 TGTGTGGGAGGAGCAACATATTACATTGTAATCTGTGCTGTGAAAGTTTGTGCAATAC 420

QY 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTGAACCGAAAGTGAAGCTTGAAGAAAGGA 480
 |||||
 Db 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTGAACCGAAAGTGAAGCTTGAAGAAAGGA 480

QY 481 GGATGGCTGTGCTTACAAAGCGCTTAAGCGAGGTGAGACCGAAATGCTGCGCAAGCT 540
 |||||
 Db 481 GGATGGCTGTGCTTACAAAGCGCTTAAGCGAGGTGAGACCGAAATGCTGCGCAAGCT 540

QY 541 GCATCCACATGACAAAGCGAAAGTGGCCAGAGCTTGCAAGTTTGTGAAGAAACAGAGAT 600
 |||||
 Db 541 GCATCCACATGACAAAGCGAAAGTGGCCAGAGCTTGCAAGTTTGTGAAGAAACAGAGAT 600

QY 601 CTCTCATAGTAATTTCTCCATGCTCAACATAGGAAAGTGTGTGCTCCCTTGGAGG 660
 |||||
 Db 601 CTCTCATAGTAATTTCTCCATGCTCAACATAGGAAAGTGTGTGCTCCCTTGGAGG 660

QY 661 TCTCTGAAGTCTCTAACCCTTGCACTCTTGGCTTCATGCTGACGACGACCTTCTAGA 720
 |||||
 Db 661 TCTCTGAAGTCTCTAACCCTTGCACTCTTGGCTTCATGCTGACGACGACCTTCTAGA 720

QY 721 TGAGCGCTTGATAGAGGATGATGACATGTTGCTGCTGAGGCTCTTGGAGAACTAG 780
 |||||
 Db 721 TGAGCGCTTGATAGAGGATGATGACATGTTGCTGCTGAGGCTCTTGGAGAACTAG 780

QY 781 AGATTTTACAGACGCTATATATAGAAAGTGTGGGAAATATAGCCAGGACTATACACA 840
 |||||
 Db 781 AGATTTTACAGACGCTATATATAGAAAGTGTGGGAAATATAGCCAGGACTATACACA 840

QY 841 TGTATCTTTCATCAATGCTTCAAGGATTTCAAGGATTTCAAGGATTTCAAGGATTTCA 900
 |||||
 Db 841 TGTATCTTTCATCAATGCTTCAAGGATTTCAAGGATTTCAAGGATTTCAAGGATTTCA 900

QY 901 ATGCACACTGAGACTAGTAACCAAGCTTCTTAAGAAAGCACTGTGCTCCATTTGCCCCC 960
 |||||
 Db 901 ATGCACACTGAGACTAGTAACCAAGCTTCTTAAGAAAGCACTGTGCTCCATTTGCCCCC 960

Db 901 ATGCACACTGAGACTAGTAACCAAGCTTCTTAAGAAAGCACTGTGCTCCATTTGCCCCC 960

QY 961 TGTCTATGCTTAGAGCTATCTGATGTCCTGAAGTGGAGAGAGTGTGTTCTGAACCTGC 1020
 |||||
 Db 961 TGTCTATGCTTAGAGCTATCTGATGTCCTGAAGTGGAGAGAGTGTGTTCTGAACCTGC 1020

QY 1021 TCTTGAATTCGTCGAAGTTCATCCAGGCGCAAGCCCTACAGCCACTTCGCAATTAAGAT 1080
 |||||
 Db 1021 TCTTGAATTCGTCGAAGTTCATCCAGGCGCAAGCCCTACAGCCACTTCGCAATTAAGAT 1080

QY 1081 GCCATACAAATGAAGCTGAGAACAGAAAGTATTCACCTGTGACCTCTGTATGCAAT 1140
 |||||
 Db 1081 GCCATACAAATGAAGCTGAGAACAGAAAGTATTCACCTGTGACCTCTGTATGCAAT 1140

QY 1141 CATCATTTGGGATCGCGAATGGCGACGCACATTAATATCCAAATCCCATTTGAACCAACT 1200
 |||||
 Db 1141 CATCATTTGGGATCGCGAATGGCGACGCACATTAATATCCAAATCCCATTTGAACCAACT 1200

QY 1201 GAAGAAAGAGAAAGATTTGGACTCAGATGCTGTCAACACCATATGAAGTCAAGTGTTC 1260
 |||||
 Db 1201 GAAGAAAGAGAAAGATTTGGACTCAGATGCTGTCAACACCATATGAAGTCAAGTGTTC 1260

QY 1261 CCCAGACTATTAACAAAGAACTTAAGGAGATCCCCAGGCGAGAAATGATCAAGAGCT 1320
 |||||
 Db 1261 CCCAGACTATTAACAAAGAACTTAAGGAGATCCCCAGGCGAGAAATGATCAAGAGCT 1320

QY 1321 GAAATCGAGCTTTAAGAGACATGTCAGTGGCTTTGGAAGAGTGTGGGATCCAGTT 1380
 |||||
 Db 1321 GAAATCGAGCTTTAAGAGACATGTCAGTGGCTTTGGAAGAGTGTGGGATCCAGTT 1380

QY 1381 CAGAGGAGGAGGATGTTGTCTCCAGTCTGGGCAAGAGATGCTATGCGGAATCTTC 1440
 |||||
 Db 1381 CAGAGGAGGAGGATGTTGTCTCCAGTCTGGGCAAGAGATGCTATGCGGAATCTTC 1440

QY 1441 TGCATAGCAGAAAGCTCCACCATTTCTTTGATGTGTTTAAAGTCTCAGCTCTC 1500
 |||||
 Db 1441 TGCATAGCAGAAAGCTCCACCATTTCTTTGATGTGTTTAAAGTCTCAGCTCTC 1500

QY 1501 TATTAATGAAGAAACAGAGCTTGTCACTCTTGTGTGCTGATGTGTAAGTATGATG 1560
 |||||
 Db 1501 TATTAATGAAGAAACAGAGCTTGTCACTCTTGTGTGCTGATGTGTAAGTATGATG 1560

QY 1561 TAGTTGAGGAAGCAATTTTCTTTGTAACCTTAAGAGTCTATTAATAAACGAGC 1620
 |||||
 Db 1561 TAGTTGAGGAAGCAATTTTCTTTGTAACCTTAAGAGTCTATTAATAAACGAGC 1620

QY 1621 ACAGATTCCACATTTTATACATGAGAGATCTTTGTGTGTAATACAGAGATTGACTGC 1680
 |||||
 Db 1621 ACAGATTCCACATTTTATACATGAGAGATCTTTGTGTGTAATACAGAGATTGACTGC 1680

QY 1681 ATCCCTTTAAAGAGTTTATGTCCCTGACTGCTGCTAAATATATCTAATTTCCAGATG 1740
 |||||
 Db 1681 ATCCCTTTAAAGAGTTTATGTCCCTGACTGCTGCTAATATATCTAATTTCCAGATG 1740

QY 1741 CTTTGTAGATGACTGAGATTTGTGAGCAATATTTGGGATTTGTAGATTGAGTGA 1800
 |||||
 Db 1741 CTTTGTAGATGACTGAGATTTGTGAGCAATATTTGGGATTTGTAGATTGAGTGA 1800

QY 1801 TGGCAGGAAAGGCGCATCTCCATTGAGATGATTAAGTGAACCAACTAGTTCTCGAAT 1860
 |||||
 Db 1801 TGGCAGGAAAGGCGCATCTCCATTGAGATGATTAAGTGAACCAACTAGTTCTCGAAT 1860

QY 1861 CTACAGAGAAAGGAGGATTCAGACTGAGAAAGTGTGACATAGACTTTAAGAACCAAAAG 1920
 |||||
 Db 1861 CTACAGAGAAAGGAGGATTCAGACTGAGAAAGTGTGACATAGACTTTAAGAACCAAAAG 1920

QY 1921 CTTTGAATTTGGAGCTGCTCATGTGTGATTAATATCACTGCTCTTTCTATTTGAGT 1980
 |||||
 Db 1921 CTTTGAATTTGGAGCTGCTCATGTGTGATTAATATCACTGCTCTTTCTATTTGAGT 1980

QY 1981 TACAAATCTATATTTTATTTGAAGTATTAATAAGAAATTTTACAGAAAAAATTTTAC 2040
 |||||
 Db 1981 TACAAATCTATATTTTATTTGAAGTATTAATAAGAAATTTTACAGAAAAAATTTTAC 2040

OY 2041 A 2041
Db 2041 A 2041

RESULT 2
AY052768 1308 bp mRNA linear PRI 17-OCT-2001
LOCUS Homo sapiens tRNA Isopentenyl transferase mRNA, partial cds.
DEFINITION AY052768
ACCESSION AY052768.1 GI:16209578
VERSION
KEYWORDS
SOURCE
ORANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1308)
AUTHORS Lemieux,J., Lakowski,B., Webb,A., Meng,Y., Ubach,A., Bussiere,F.,
Barnes,T. and Hekimi,S.
TITLE Regulation of physiological rates in Caenorhabditis elegans by a
tRNA-modifying enzyme in the mitochondria
JOURNAL Genetics 159 (1), 147-157 (2001)
MEDLINE 21444833
PUBMED 11560893
REFERENCE 2 (bases 1 to 1308)
AUTHORS Lemieux,J., Barnes,T. and Hekimi,S.
TITLE Direct Submission
JOURNAL Submitted (28-AUG-2001) Biology, McGill University, 1205 Dr.
Penfield Avenue, Montreal, QC H3A 1B1, Canada
FEATURES
source
1. 1308
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-2ec05"
<1. 1308
/note="hgr01"
/codon_start=1
/product="tRNA isopentenyl transferase"
/protein_id="AAL14107.1"
/db_xref="GI:16209579"
/translacion="AARVPGSGGLGRLRLPLVVIIGANGTGKSTLALQLGRLGGE
IVSADSMOVEGLDITNKVSAOERICRHMISFVDPVLTNYVDFRNATLIED
IFARDKPIVNGNYIESILKLVNTKOEKTEKVIDKYLEDEGLVLAHKL
SOVPEMAKAKHPHDKRVARSLQVPEFTGSHSEPLRROTREBGGGLGSLPLFKNP
CITLHADQAVLDERLDKRVDDMLAAGLELRDFHRYRNKNVSENSODOHGFOS
IGFEFEHYLTTECKTLETSLNQLKPGPIVPPVGLVSDVSKMESVLEPALEI
VOSFIQGHKPTATPDKMPYNEAKNRKSYHLCDLDRILIGREVAAHIKSKSHNOLK
KRRLDSPAVMTIESOSVSPDYNEKPEKSGPGQNDLKCYSV"

BASE COUNT 378 a 302 c 329 g 299 t

ORIGIN

Query Match 61.6%; Score 1258; DB 9; Length 1308;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 GCACGACAGTTCTGTGGGCGAGTGGGCTCAGGGCGCTGCAACGACCTTACTCTTGA 88
Db 1 GCACGACAGTTCTGTGGGCGAGTGGGCTCAGGGCGCTGCAACGACCTTACTCTTGA 60

OY 89 GTGATTCGCGGGCGACGCGGCGCAATCAGCGTGGGCTGACGTAGCGGCGG 148
Db 61 GTGATTCGCGGGCGACGCGGCGCAATCAGCGTGGGCTGACGTAGCGGCGG 120

OY 149 CTCGGCGGTGAGATCGTCAGCGCTGACTGCATGCAAGGTCTATGAAGCGCTAGACATC 208
Db 121 CTCGGCGGTGAGATCGTCAGCGCTGACTGCATGCAAGGTCTATGAAGCGCTAGACATC 180

OY 209 ACCAACAGGTTTCTGCCCCAGACAGAGATCTGCGGCGACCATGATCATGCTTTGTG 268
Db 181 ACCAACAGGTTTCTGCCCCAGACAGAGATCTGCGGCGACCATGATCATGCTTTGTG 240

OY 269 GATCTCTTGTGACCAATTACACAGTGGTGCATTCAGAAATAGAGCACTCTCTGATT 328

Db 241 GATCCTCTTGTGACCAATTACACAGTGGTGCATTCAGAAATAGAGCACTCTCTGATT 300
OY 329 GAAGTATATTTGCCCCGACACAAATTCCTATTGTTGGGAGGAGACCAATTATATTAATT 388
Db 301 GAAGTATATTTGCCCCGACACAAATTCCTATTGTTGGGAGGAGACCAATTATATTAATT 360

OY 389 GAATCTCTGCTCTGGAAGATTCTTGTCATACCAAGCCCGAGAGANTGGGCACTAGAAA 448
Db 361 GAATCTCTGCTCTGGAAGATTCTTGTCATACCAAGCCCGAGAGANTGGGCACTAGAAA 420

OY 449 GTGATTGACCGAAAGAGTGGACCTTGAAAAGAGATGCTTTGATTCTTACAAACGCTTA 508
Db 421 GTGATTGACCGAAAGAGTGGACCTTGAAAAGAGATGCTTTGATTCTTACAAACGCTTA 480

OY 509 AGCCAGTGGAGCCCAAGAAATGGCTGCCATTCACATGACAAACGAAAGTGGCC 568
Db 481 AGCCAGTGGAGCCCAAGAAATGGCTGCCATTCACATGACAAACGAAAGTGGCC 540

OY 569 AGGACCTTGCAAGTTTGAAGAAACAGAAATCTCTCATAGTAATTTCTCCATGCTCAA 628
Db 541 AGGACCTTGCAAGTTTGAAGAAACAGAAATCTCTCATAGTAATTTCTCCATGCTCAA 600

OY 629 CATACGGAAGAGTGTGTCCTCCCTTGAGAGTCTCTGAAGTCTTAACCTTGATC 688
Db 601 CATACGGAAGAGTGTGTCCTCCCTTGAGAGTCTCTGAAGTCTTAACCTTGATC 660

OY 689 CTTTGGCTTCATGCTGACAGGACCTTCTAGATGAGCGCTTGATTAAGAGGTGATGAC 748
Db 661 CTTTGGCTTCATGCTGACAGGACCTTCTAGATGAGCGCTTGATTAAGAGGTGATGAC 720

OY 749 ATGCTTGCTGCTGGGCTCTTGAGAGAACTAAGATTTTTCACAGAGCTATTAATGAGAAG 808
Db 721 ATGCTTGCTGCTGGGCTCTTGAGAGAACTAAGATTTTTCACAGAGCTATTAATGAGAAG 780

OY 809 AATGTTTGGAAATAGGCAGAGATATCAATGATGTTCTTCAATTAATGCTTCAAG 868
Db 781 AATGTTTGGAAATAGGCAGAGATATCAATGATGTTCTTCAATTAATGCTTCAAG 840

OY 869 GAATTTTCAGAGTACTGATCTACTGAGGAAATATGACACTGAGAACTAGTAAACAGCTT 928
Db 841 GAATTTTCAGAGTACTGATCTACTGAGGAAATATGACACTGAGAACTAGTAAACAGCTT 900

OY 929 CTAAGAAGAGACCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTC 988
Db 901 CTAAGAAGAGACCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTC 960

OY 989 TCGAAGTGGAGAGTCTGTTCTTGAACCTGCTTGAATGTCGAAGTTTCATCCAG 1048
Db 961 TCGAAGTGGAGAGTCTGTTCTTGAACCTGCTTGAATGTCGAAGTTTCATCCAG 1020

OY 1049 GGCCACAAAGCTTACAGCCACTCCATTAAGATGCCATATGAAGCTGACACAAAGAGA 1108
Db 1021 GGCCACAAAGCTTACAGCCACTCCATTAAGATGCCATATGAAGCTGACACAAAGAGA 1080

OY 1109 AATTATACCTGCTGACCTGTCCTGATGCAATCATTTGGGATCGGGAATGGGCGAGG 1168
Db 1081 AATTATACCTGCTGACCTGTCCTGATGCAATCATTTGGGATCGGGAATGGGCGAGG 1140

OY 1169 CACATTAATCCCAATTCACACTTGAACCACTGAAGAAAGAGAAGATTGGACTCAGAT 1228
Db 1141 CACATTAATCCCAATTCACACTTGAACCACTGAAGAAAGAGAAGATTGGACTCAGAT 1200

OY 1229 GCTGTCAACACCATGAAGAGTGTGTTTCCCAAGACTATTAACAAAGACCTTAAG 1286
Db 1201 GCTGTCAACACCATGAAGAGTGTGTTTCCCAAGACTATTAACAAAGACCTTAAG 1258

RESULT 3
AX405704 2130 bp DNA linear PAT 14-JUN-2002
LOCUS AX405704
DEFINITION Sequence 119 from Patent WO0222660.
ACCESSION AX405704


```

VERSION      AX405704.1  GI:21438847
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
REFERENCE   1
AUTHORS     Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F.,
            Xue, A. J., Yang, Y., Mehrman, T. and Drmanac, R. T.
TITLE       Novel nucleic acids and polypeptides
JOURNAL     Patent: WO 0222660-A 119 21-MAR-2002;
            HYSEQ, INC. (US)
FEATURES
source      1..2130
            Location/Qualifiers
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            191..1426
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD34811.1"
            /db_xref="GI:21438848"
            /translation="MOVYEGLDITNKVSAQEDRICRHHMISFVDPVLYNRYTVVDFRN
            KATALIEDIFARDKLPVIVGGTNYTIESLWKVLVNTKPOEMGTEKVIDKVELKED
            GLVLRHSQVDPENAAKLPHDKRKVARSLOVEETGISHEFLRHOTEEGGPLG
            GPLKFSNPCLMLHADAVALDERLDKRVDDMLAAGLEELDFHRRYNQKNVSENSOD
            YOHGISIGKEFEHEYLITEGKCTLETNSDLKGLIEALQVTKRYARKONRWKRN
            FLSRGPVLPVYGLEISDVSKWESVLEPALIEIYOSFIOGHKPTAPPIKMPYNAEN
            KSYHLCDLCDRITIGDREMAAHIKSKSHLNQKKRRRLDSADVNTIESOSVSPDHKK
            EPEKSPQNDDELKCSV"

BASE COUNT  620 a 439 c 522 g 549 t
ORIGIN
Query Match      43.9% Score 896; DB 6; Length 2130;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 896; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  44 GTGGCAGTGGGCTCAGAGGCTCTGCAAGACCTTACCTCTTGTAGTATTCGCGGCC 103
    |||||
DB  56 GTGGCAGTGGGCTCAGAGGCTCTGCAAGACCTTACCTCTTGTAGTATTCGCGGCC 115
    |||||

OY  104 ACGGCGACGGCAATCCAGCGTGGCTTCAGCTAGGCCAGCGGCTCGCGGTGACATC 163
    |||||
DB  116 ACGGCGACGGCAATCCAGCGTGGCTTCAGCTAGGCCAGCGGCTCGCGGTGACATC 175
    |||||

OY  164 GTACGGCTGACGTCATGACAGGTCTATGAAGGCTTAGACATCATCACAAGGCTTCT 223
    |||||
DB  176 GTACGGCTGACGTCATGACAGGTCTATGAAGGCTTAGACATCATCACAAGGCTTCT 235
    |||||

OY  224 GCCCAAGAGCAGAGAAATCTCCGCGACCATGATCAGCTTGTGATCCTCTTGTGACC 283
    |||||
DB  236 GCCCAAGAGCAGAGAAATCTCCGCGACCATGATCAGCTTGTGATCCTCTTGTGACC 295
    |||||

OY  284 AATTACAGTGTGTGACCTCAGAAATAGACAGCACTGCTGATTTGAAGATATTTGGC 343
    |||||
DB  296 AATTACAGTGTGTGACCTCAGAAATAGACAGCACTGCTGATTTGAAGATATTTGGC 355
    |||||

OY  344 CGAGACAAATTCCTATTTGTGGGAGGAACAATATTAATGAATCTGCTGCTGG 403
    |||||
DB  356 CGAGACAAATTCCTATTTGTGGGAGGAACAATATTAATGAATCTGCTGCTGG 415
    |||||

OY  404 AAAGTTCTTGTCAATACCAAGCCCCAGAGAGTGGGCACTGAGAAAGTATTGACGAAAA 463
    |||||
DB  416 AAAGTTCTTGTCAATACCAAGCCCCAGAGAGTGGGCACTGAGAAAGTATTGACGAAAA 475
    |||||

OY  464 GTGGAGCTTGAAGAGAGATGCTTTGTACTTACAAACGCTAAGCCAGGTGAGACCA 523
    |||||
DB  476 GTGGAGCTTGAAGAGAGATGCTTTGTACTTACAAACGCTAAGCCAGGTGAGACCA 535
    |||||

OY  524 GAATGCTGCCCAAGCTGCATCCACATGACAAACGCAAAATGGCCAGGACTTGCAGATT 583
    |||||
DB  536 GAATGCTGCCCAAGCTGCATCCACATGACAAACGCAAAATGGCCAGGACTTGCAGATT 595
    |||||

OY  584 TTGAAGAAACAGGAATCTCTCATAGTAATTTCTCATGCTGATACATAGCAAGAGT 643

```

```

DB  596 TTGAAGAAACAGGAATCTCTCATAGTAATTTCTCATGCTGATACATAGCAAGAGT 655
    |||||

OY  644 GGTGTCCTCTTGGAGGCTCTCTGAACTTCTTAACCTTGCATCTTTGGCTTCATGCT 703
    |||||
DB  656 GGTGTCCTCTTGGAGGCTCTCTGAACTTCTTAACCTTGCATCTTTGGCTTCATGCT 715
    |||||

OY  704 GACGAGCACTTCTAGATGAGCGCTTGATTAAGGGGTGATGATGATGCTTGGCTGGG 763
    |||||
DB  716 GACGAGCACTTCTAGATGAGCGCTTGATTAAGGGGTGATGATGATGCTTGGCTGGG 775
    |||||

OY  764 CTCTTGAGAGAACTAAGAGATTTTTCACAGACGCTATTAATCAGAAATTTTGGAAAT 823
    |||||
DB  776 CTCTTGAGAGAACTAAGAGATTTTTCACAGACGCTATTAATCAGAAATTTTGGAAAT 835
    |||||

OY  824 AGCCAGACATATCAACATGATCTTCAATCAATTTGGCTTCAAGAAATTTACAGACTAC 883
    |||||
DB  836 AGCCAGACATATCAACATGATCTTCAATCAATTTGGCTTCAAGAAATTTACAGACTAC 895
    |||||

OY  884 CTGATCAGTGAAGGAAATCAGACCTGAGACATAGTAACAGCTTCTTAAGAAAG 939
    |||||
DB  896 CTGATCAGTGAAGGAAATCAGACCTGAGACATAGTAACAGCTTCTTAAGAAAG 951
    |||||

RESULT 4
AF074918
LOCUS
DEFINITION Homo sapiens tRNA isopentenylpyrophosphate transferase precursor
            RNA, complete cds.
VERSION     AF074918
KEYWORDS    AF074918.1 GI:11137964
ORGANISM    Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
REFERENCE   1 (bases 1 to 2216)
AUTHORS     Golovko, A., Hjalim, G., Stedon, F. and Nicander, B.
TITLE       Cloning of a human tRNA isopentenyl transferase
JOURNAL     MEDLINE 20564178
            PUBMED 11111046
REFERENCE   2 (bases 1 to 2216)
AUTHORS     Golovko, A. and Hjalim, G.
TITLE       Direct Submission
JOURNAL     Submitted (28-JUN-1998) Dep. of Plant Biology, Swedish University
            of Agricultural Sciences, Box 7080, SE-750 07 Uppsala, Sweden
FEATURES
source      1..2216
            Location/Qualifiers
            /organism="Homo sapiens"
            /db_xref="dbEST:AA332152"
            /db_xref="taxon:9606"
            /note="extension of sequence by 5' RACE"
            1..2216
            /note="incompletely processed mRNA"
            join(15..829,917..1505)
            /note="1Pase"
            /codon_start=1
            /product="tRNA isopentenylpyrophosphate transferase"
            /protein_id="AA331324.1"
            /db_xref="GI:11137965"
            /translation="MASVAARAARVPGSLRGLRPLPLVILLAGTGKSTLALQIG
            ORLGEIVASAGMOYEGDITITNKVSAQEDRICRHHMISFVDPVLYNRYTVVDFRRA
            TALIEDIFARDKLPVIVGGTNYTIESLWKVLVNTKPOEMGTEKVIDRVELEKEDGL
            VLHRLSQVDPENAAKLPHDKRKVARSLOVEETGISHEFLRHOTEEGGPLGGLKFSNPCL
            MLHADAVALDERLDKRVDDMLAAGLEELDFHRRYNQKNVSENSODYOHGISIGKEFEHE
            YLITEGKCTLETNSDLKGLIEALQVTKRYARKONRWKRNFLSRGPVLPVYGLEISDVSKW
            ESVLEPALIEIYOSFIOGHKPTAPPIKMPYNAENKSYHLCDLCDRITIGDREMAAHIKSK
            SHLNQKKRRRLDSADVNTIESOSVSPDHKKERKSPQNDDELKCSV"
            830..916
            /evidence="experimental"
            2177..2182
            polyA_signal

```



```

Db 983 CCAATCCACCTGAAACCACTGAAGAAAAGAAAGATTGGACTGAGTGCCTGACACA 1042
Qy 1239 CCATTAAGAGTCAGAGTGTTCCTCCAGACTATATACAAAGAACTAAAGGAGGATCCC 1298
Db 1043 CCATTAAGAGTCAGAGTGTTCCTCCAGACTATATACAAAGAACTAAAGGAGGATCCC 1102
Qy 1299 CAGGCGAGATGATCAAGAGCTGAATGACGCTTTAAGAGCATGCTCCAGTGGCCTTG 1358
Db 1103 CAGGCGAGATGATCAAGAGCTGAATGACGCTTTAAGAGCATGCTCCAGTGGCCTTG 1162
Qy 1359 GAAAGTGTGTGGGATCCAGATTGAGAGGAGGAGGATATGTTGTCTCCAGTGGGCAA 1418
Db 1163 GAAAGTGTGTGGGATCCAGATTGAGAGGAGGAGGATATGTTGTCTCCAGTGGGCAA 1222
Qy 1419 AGGAGTGTGTGGGATCCAGATTGAGAGGAGGAGGATATGTTGTCTCCAGTGGGCAA 1478
Db 1223 AGGAGTGTGTGGGATCCAGATTGAGAGGAGGAGGATATGTTGTCTCCAGTGGGCAA 1282
Qy 1479 GGTTTAAAGTGTGAGTGTCTCTATATATGAAGACAGAGTGTCTGACGCTCTGTCG 1538
Db 1283 GGTTTAAAGTGTGAGTGTCTCTATATATGAAGACAGAGTGTCTGACGCTCTGTCG 1342
Qy 1539 GCTGATGTGTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1598
Db 1343 GCTGATGTGTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1402
Qy 1599 AGTGTCTATATATATATATATATATATATATATATATATATATATATATATATAT 1658
Db 1403 AGTGTCTATATATATATATATATATATATATATATATATATATATATATATATAT 1462
Qy 1659 GGTGAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1718
Db 1463 GGTGAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1521
Qy 1719 AAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 1778
Db 1522 AAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 1581
Qy 1779 GGGAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1838
Db 1582 GGGAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1641
Qy 1839 AACCAACTAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1898
Db 1642 AACCAACTAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1701
Qy 1899 CATAGACTTGAAGACCAAGACTTGAATTTGAGCTGCTCATGCTGATGATGATGAT 1958
Db 1702 CATAGACTTGAAGACCAAGACTTGAATTTGAGCTGCTCATGCTGATGATGATGAT 1761
Qy 1959 CACTGCTGTCTTCTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2018
Db 1762 CACTGCTGTCTTCTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1821
Qy 2019 AAATTACAA 2028
Db 1822 AAATTACAA 1831

```

```

RESULT 6
AK000068 1749 bp mRNA linear PRI 22-FEB-2000
DEFINITION Homo sapiens CDNA FLJ20061 fls. clone COL01383.
ACCESSION AK000068
VERSION AK000068.1 GI:7019914
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens colon CDNA to mRNA, clone_11b:COL clone:COL01383.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,

```

```

FEATURES
source
    Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
    Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
    MEDO human CDNA sequencing project
    Unpublished
    2 (bases 1 to 1749)
    Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
    Shibahara,T., Tanaka,T. and Nakamura,Y.
    Direct Submission
    Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
    University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
    Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp,
    Tel:81-3-5449-5286, Fax:81-3-5449-5416)
    MEDO human CDNA sequencing project supported by Ministry of
    International Trade and Industry of Japan; CDNA full insert
    sequencing; Research Association for Biotechnology; CDNA library
    construction, 5'-3'-end one pass sequencing; Department of
    Virology and Human Genome Center, Institute of Medical Science,
    University of Tokyo (partly supported by Science and Technology
    Agency).

    1..1749
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="COL01383"
    /tissue_type="colon"
    /clone_11b="COL"
    /note="cloning vector pME185FL3"
    61..1041
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="BAA09023.1"
    /db_xref="GI:7019915"
    /translation="MTEKEVVIDRKVELEKEDGLVLRKLSQVDPMAKLRHPRDKRY
    ARSLQVFEETGISHSEFLRHQHTTEGGGRLGFLFNSPCLIMLAADAVLDERDKR
    VDDMLAAGLIEELRDFHRRYNOKNVSENSODHGIPOSIGFKEHEVLITGKCTLE
    TSNOLKKGIEALKQVTRKARKONRWKRNRLSPGPIPVYGLVSDSKMESV
    LEPALEIVOSFIOGHKPTATPKMYNRYENKRSYHLIDICRIIIGREMAAHLKSK
    SHLNOLKRRRLDSQAVNTIESQSVSPDHNNKPKERKSGQNDDELKCSV"

BASE COUNT 541 a 337 c 408 g 463 t
ORIGIN
Query Match 36.8%; Score 751; DB 9; Length 1749;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1021; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

Oy 1421 GAGTGTATGCGGAAATTCCTGATAGCAGAAAAGCTCCACATTTCTTTTGATGTGC 1480
 |||||||
 Db 1126 GAGGCGATATCGGAAATTCCTGATAGCAGAAAAGCTCCACATTTCTTTGATGTGC 1185
 |||||||
 Oy 1481 TTTTAAAGCTCAGCTTCTCTATATAGAAAACAGCAGCTGTGTGCTCTGTGTGC 1540
 |||||||
 Db 1186 TTTTAAAGCTCAGCTTCTCTATATAGAAAACAGCAGCTGTGTGCTCTGTGTGC 1245
 |||||||
 Oy 1541 TGATGCTCTGGAATATGATAGTTCAGAAAAGCATTTTCTTTTGACCTTAAG 1600
 |||||||
 Db 1246 TGATGCTCTGGAATATGATAGTTCAGAAAAGCATTTTCTTTTGACCTTAAG 1305
 |||||||
 Oy 1601 GTTCTATTATTAAGCAGACAGATTCACATTTTATACATGAGATCTTTGTGTGC 1660
 |||||||
 Db 1306 GTTCTATTATTAAGCAGACAGATTCACATTTTATACATGAGATCTTTGTGTGC 1365
 |||||||
 Oy 1661 TGAATACAGAGATTGACTGCATCCCTTTAAAGAGTTTATGTCCCTGACTGTGCTAA 1720
 |||||||
 Db 1366 TGAATACAGAGATTGACTGCATCCCTTTT -AAAGAGTTTATGTCCCTGACTGTGCTAA 1424
 |||||||
 Oy 1721 AATTATCTAATTTCCAGATGCTTTTGTAGATGACTGAAGTATTTGTGACCAATATTCG 1780
 |||||||
 Db 1425 AATTATCTAATTTCCAGATGCTTTTGTAGATGACTGAAGTATTTGTGACCAATATTCG 1484
 |||||||
 Oy 1781 GAGTTCTAGATTTGAGTGAATGCGAGAAAGGCCATCTCCATGAGATGATTAAGTGA 1840
 |||||||
 Db 1485 GAGTTCTAGATTTGAGTGAATGCGAGAAAGGCCATCTCCATGAGATGATTAAGTGA 1544
 |||||||
 Oy 1841 CCAAACTAGATTCTCGAATTCATACAGAAAGGAGGAATTCAGACTGAGAAAGCTGTGACA 1900
 |||||||
 Db 1545 CCAAACTAGATTCTCGAATTCATACAGAAAGGAGGAATTCAGACTGAGAAAGCTGTGACA 1604
 |||||||
 Oy 1901 TAGACTTGAAGACCAAAAGACTTTGAAATTTGCGAGCTGATGTGACTTATTAACA 1960
 |||||||
 Db 1605 TAGACTTGAAGACCAAAAGACTTTGAAATTTGCGAGCTGATGTGACTTATTAACA 1664
 |||||||
 Oy 1961 CTGCTGCTTCTCTATGAGTACAAATCTATATTTTATTAAGTTTAAATTAAGAAAAA 2020
 |||||||
 Db 1665 CTGCTGCTTCTCTATGAGTACAAATCTATATTTTATTAAGTTTAAATTAAGAAAAA 1724
 |||||||
 Oy 2021 ATTTA 2025
 |||||||
 Db 1725 ATTTA 1729
 |||||||

RESULT 7
 AK074222
 LOCUS
 DEFINITION
 AK074222 1749 bp mRNA linear PRI 15-FEB-2002
 Homo sapiens cDNA FLJ23642 fis, clone COL01371, highly similar to
 Homo sapiens tRNA isopentenylpyrophosphate transferase precursor
 RNA.
 ACCESSION
 AK074222 GI:18676765
 VERSION
 AK074222.1 GI:18676765
 KEYWORDS
 oligo capping; fis (full insert sequence).
 SOURCE
 Homo sapiens colon cDNA to mRNA, clone_11b.COL clone:COL01371.
 ORGANISM
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 1
 Kawabata,A., Hiki,T., Kobatake,N., Inagaki,H., Ikema,Y.,
 Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
 Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project
 TITLE
 JOURNAL
 SUGANO,S., SUZUKI,Y., OTA,T., ODAYASHI,M., NISHI,T., ISOGAI,T.,
 SHIBAHARA,T., TANAKA,T. and NAKAMURA,Y.
 REFERENCE
 AUTHORS
 2 (bases 1 to 1749)
 Direct Submission
 NEDO human cDNA sequencing project
 TITLE
 JOURNAL
 Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure, Human Genome
 Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
 (E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)

COMMENT
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology; cDNA library construction,
 5' - 6' 3'-end one pass sequencing: Department of Virology and Human
 Genome Center, Institute of Medical Science, University of Tokyo
 (partly supported by Science and Technology Agency).
 FEATURES
 source
 1. 1749
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="COL01371"
 /tissue_type="colon"
 /clone_11b="COL"
 /note="cloning vector: pME18SFL3"
 BASE COUNT 541 a 337 c 408 g 463 t
 ORIGIN
 Query Match 36.8%; Score 751; DB 9; Length 1749;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1021; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 Oy 1001 GAGTGTATGCGGAAATTCCTGATAGCAGAAAAGCTCCACATTTCTTTTGATGTGC 1060
 |||||||
 Db 706 GAGTGTATGCGGAAATTCCTGATAGCAGAAAAGCTCCACATTTCTTTGATGTGC 765
 |||||||
 Oy 1061 ACAGCCACTCCAAATAAGATGCAATACATGAGCTGAGAACAGAGAAATTCACCTG 1120
 |||||||
 Db 766 ACAGCCACTCCAAATAAGATGCAATACATGAGCTGAGAACAGAGAAATTCACCTG 825
 |||||||
 Oy 1121 TGTGACCTCTGTGATCGAATCATCATTTGGGATGCGCAATGGCAGCCACATTAATTC 1180
 |||||||
 Db 826 TGTGACCTCTGTGATCGAATCATCATTTGGGATGCGCAATGGCAGCCACATTAATTC 885
 |||||||
 Oy 1181 AAATCCCACTGAACCACTGAAGAAAAGAAAGATTTGAGTGCATGATGCTGTGAACGC 1240
 |||||||
 Db 886 AAATCCCACTGAACCACTGAAGAAAAGAAAGATTTGAGTGCATGATGCTGTGAACGC 945
 |||||||
 Oy 1241 ATAGAAGTCAAGAGTGTTCCTCCAGACTATTAACAAGAACTTAAAGGAGGATCCCA 1300
 |||||||
 Db 946 ATAGAAGTCAAGAGTGTTCCTCCAGACTATTAACAAGAACTTAAAGGAGGATCCCA 1005
 |||||||
 Oy 1301 GGGCAGATGATCAAGAGCTGAAGTGCAGCTTTTAAAGACATGTCCAGTGGCTTTGGA 1360
 |||||||
 Db 1006 GGGCAGATGATCAAGAGCTGAAGTGCAGCTTTTAAAGACATGTCCAGTGGCTTTGGA 1065
 |||||||
 Oy 1361 AAGGTGGGGGATCCAGTTCAGAGGAGGAGGATGTGTTGTCCTCCAGTGGGCAAG 1420
 |||||||
 Db 1066 AAGGTGGGGGATCCAGTTCAGAGGAGGAGGATGTGTTGTCCTCCAGTGGGCAAG 1125
 |||||||
 Oy 1421 GAGTGTATGCGGAAATTCCTGATAGCAGAAAAGCTCCACATTTCTTTTGATGTGC 1480
 |||||||
 Db 1126 GAGTGTATGCGGAAATTCCTGATAGCAGAAAAGCTCCACATTTCTTTTGATGTGC 1185
 |||||||
 Oy 1481 TTTTAAAGCTCAGCTTCTCTATATAGAAAACAGCAGCTCTTGTGAGTGC 1540
 |||||||
 Db 1186 TTTTAAAGCTCAGCTTCTCTATATAGAAAACAGCAGCTCTTGTGAGTGC 1245
 |||||||
 Oy 1541 TGATGCTCTGGAATATGATAGTTCAGAAAAGCATTTTCTTTTGACCTTAAG 1600
 |||||||
 Db 1246 TGATGCTCTGGAATATGATAGTTCAGAAAAGCATTTTCTTTTGACCTTAAG 1305
 |||||||
 Oy 1601 GTTCTATTATTAAGCAGACAGATTCACATTTTATACATGAGATCTTTGTGTGC 1660
 |||||||
 Db 1306 GTTCTATTATTAAGCAGACAGATTCACATTTTATACATGAGATCTTTGTGTGC 1365
 |||||||
 Oy 1661 TGAATACAGAGATTGACTGCATCCCTTTAAAGAGTTTATGTCCCTGACTGTGCTAA 1720
 |||||||
 Db 1366 TGAATACAGAGATTGACTGCATCCCTTT -AAAGAGTTTATGTCCCTGACTGTGCTAA 1424
 |||||||
 Oy 1721 AATTATCTAATTTCCAGATGCTTTTGTAGATGACTGAAGTATTTGTGAGCCACATATTGC 1780
 |||||||
 Db 1425 AATTATCTAATTTCCAGATGCTTTTGTAGATGACTGAAGTATTTGTGAGCCACATATTGC 1484
 |||||||

OY 1781 GAGTCTAGATTGAGTGAATGCGAAGAGGCCATCTCCATTGAGATGATTAACTGAA 1840
|||||
DB 1485 GAGTCTAGATTGAGTGAATGCGAAGAGGCCATCTCCATTGAGATGATTAACTGAA 1544
|||||
OY 1841 CCAACTAGTCTCGAATTTCTACAGAGAAGGAGGAGTACAGCTGAGAGCTGTGACA 1900
|||||
DB 1545 CCAACTAGTCTCGAATTTCTACAGAGAAGGAGGAGTACAGCTGAGAGCTGTGACA 1604
|||||
OY 1901 TAGACTGTGAGACCAAGACTTGAATTTGCGAGCTGCTCATGTGTGATTAATCA 1960
|||||
DB 1605 TAGACTGTGAGACCAAGACTTGAATTTGCGAGCTGCTCATGTGTGATTAATCA 1664
|||||
OY 1961 CTGCTCTCTTCTATTGAGTTACAACTATATTTTATGAGTTAAATAAGAAAAA 2020
|||||
DB 1665 CTGCTCTCTTCTATTGAGTTACAACTATATTTTATGAGTTAAATAAGAAAAA 1724
|||||
OY 2021 ATTTA 2025
|||||
DB 1725 ATTTA 1729
|||||
RESULT 8
HS118J21/c 166496 bp DNA linear PRI 18-FEB-2000
LOCUS
DEFINITION Human DNA sequence from clone RPI-118J21 on chromosome 1p34.1-35.3
Contains part of the gene for BMP8 (bone morphogenetic protein 8
(osteogenic protein 2)), an l-myc-proto-oncogene, STS, GSSS and
Cpg Islands.n, complete sequence.
ACCESSION AL033527
VERSION AL033527.26 GI:6456853
KEYWORDS HTG: BMP8; Cpg Island; l-myc; morphogenetic; oncogene; osteogenic.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 166496)
E. Blington, A.
Direct Submission
Submitted (18-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Nov 21, 1999 this sequence version replaced gi:6002299.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL, Sw: SWISSPROT, Tr: TREMBL, Wp: WORMPEP, Infomation
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/c_elegans/wormpep RPI-118J21 is
from the library RPI-1 constructed at the Roswell Park Cancer
Institute by the group of Pieter de Jong. For further details see
http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
This sequence is the entire insert of clone RPI-118J21 The true
right end of clone RPI-117L23 is at 38518 in this sequence. This
sequence was generated from part of bacterial clone contigs of
human chromosome 1, constructed by the Sanger Centre Chromosome 1
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1.
FEATURES
source
1..166496
/organism="Homo sapiens"
/db_xref="taxon:9606"
Location/Qualifiers

/chromosome="1"
/map="p34.1-35.3"
/clone="RPI-118J21"
/clone_11b="RPI-1"
complement(830..29381)
/gene="BMP8"
complement(join(<830..1098,3622..3732,4242..4321,
5153..5347,14838..14986,15379..15568,26862..29381))
/gene="BMP8"
/product="d1118J21.1 (bone morphogenetic protein 8
(osteogenic protein 2))"
/note="match: CDNAS: Em:577477 Em:X51801 Em:M97017
Em:U39545 Em:J04566 Em:X80992 Em:M60316 Em:M60315
Em:M97016 Em:X56906 Em:AF100787 Em:U40034"
/evidence=not_experimental
complement(join(949..1098,3622..3732,4242..4321,
5153..5347,14838..14986,15379..15568,26862..29015))
/gene="BMP8"
/note="match: proteins: Sw:P22004.Sw:P20722 Sw:P22003
Sw:P18075 Tr:Q9YGH7 Sw:P23359 Sw:P55105 Sw:P49003
Sw:P34820 Sw:P34821"
/codon_start=1
/evidence=not_experimental
/product="d1118J21.1 (bone morphogenetic protein 8
(osteogenic protein 2))"
/protein_id="CB75681.1"
/db_xref="GI:7018292"
/translation="MTALPGPLMLGLALCALGCGGPGRLRPPGCPORRLGARRRDY
OREILAVLGLPGRPRPAPASRLPASPFLPMDLYHAMAGDDDEDCAFERRGRRA
DLVMSFVNVERDRALGHOEPHMKFERFDLQIPAGEAVTAEFRIYVPSHILNRT
LHVSFMOVVOEGSNRESDFLDLQTLRAGDCGMLVDTAASDCMLLKRHKDGLRL
VYETEDGHSVDPGLGLGQRPASQGPVYVTFRRASPSPIITPRAYPLRLRROKKS
NELPQANRLPGITFDVGHSGHGRQVCRHRELVYFSDGLMDLVNAPQSYSAVYCEGC
SEPLDSCKNATNHAIIQSLVHLMMPDAVPAKCAFPKLSATSVLYSDSNVILKRR
NMVAVACCH"
1655..1748
/note="L2 repeat: matches 1691..1787 of consensus"
1849..2281
/note="L2 repeat: matches 2002..2438 of consensus"
2764..2868
/note="FLAM_C repeat: matches 1..133 of consensus"
complement(4214..4420)
/gene="BMP8"
/note="match: STS: Em:T63436"
complement(4252..4438)
/gene="BMP8"
/note="match: STS: Em:AA634039"
5972..6029
/note="MIR repeat: matches 19..75 of consensus"
6293..6588
/note="Alub repeat: matches 1..307 of consensus"
7787..7873
/note="MIR repeat: matches 57..146 of consensus"
complement(8369..8876)
/gene="BMP8"
/note="match: GSS: Em:B51067"
9109..9404
/note="Alub repeat: matches 1..296 of consensus"
9427..9746
/note="Alub repeat: matches 1..312 of consensus"
10075..10242
/note="match: STS: Em:Z41229"
10626..12048
/note="Cpg Island"
/evidence=not_experimental
complement(L2 repeat: matches 957..2141 of consensus"
12351..13402
/note="L2 repeat: matches 178..295 of consensus"
13474..13591
/note="Aluo/FRAM repeat: matches 178..295 of consensus"
13596..14185
/note="L1MC3 repeat: matches 7071..7691 of consensus"
14186..14425
/note="L2 repeat: matches 2129..2417 of consensus"

Query Match	Best Local Similarity	29.1%	Score 593;	DB 9;	Length 166496;
Matches 863;	Conservative	99.5%;	Pred. No. 5.7e-313;		
		0;	Mismatches 3;	Indels 1;	Gaps 1;
Repeat_Region	33075..33487				
Repeat_Region	/note="L2 repeat: matches 2278..2701 of consensus"				
Repeat_Region	33804..34154				
Repeat_Region	/note="Alusx repeat: matches 3..294 of consensus"				
QY	1164	CAGCCGACATATAATTCACAAATCCACCTTGAAACCACTAGAGAAAAGAAAGAGATTGGACT	1223		
DB	82446	CAGCCGACATATAATTCACAAATCCACCTTGAAACCACTAGAGAAAAGAAAGATTGGACT	82387		
QY	1224	CAGATGCTGTCAACACCATAGAAATCTCAGAGTGTTCCTCCAGACTATTAACAAGAACCTA	1283		
DB	82386	CAGATGCTGTCAACACCATAGAAATCTCAGAGTGTTCCTCCAGACTATTAACAAGAACCTA	82327		
QY	1284	AAGGAGAGGAGATCCCGACGGCGAGATGATCAAGACCTAAATGACGCGCTTTAAAGACAT	1343		
DB	82326	AAGGAGAGGAGATCCCGACGGCGAGATGATCAAGACCTAAATGACGCGCTTTAAAGACAT	82267		
QY	1344	GTCAGATGCGCTTTGGAAAGGTGTGGGGATCCACTCAGAGGAGGGGTATGTTTGTCT	1403		
DB	82266	GTCAGATGCGCTTTGGAAAGGTGTGGGGATCCACTCAGAGGAGGGGTATGTTTGTCT	82207		
QY	1404	TCCCACTGTGGGCAAAAGAGTCTATGCGGATTTCTCTGCATAGCAGAAAACCTCCACCC	1463		
DB	82206	TCCCACTGTGGGCAAAAGAGTCTATGCGGATTTCTCTGCATAGCAGAAAACCTCCACCC	82147		
QY	1464	ATTTTCCTTTGATGTGGTTTAAAGCTCAGCTCTCTTAATATAGCAAGACGCTCTTG	1523		
DB	82146	ATTTTCCTTTGATGTGGTTTAAAGCTCAGCTCTCTTAATATAGCAAGACGCTCTTG	82087		
QY	1524	TCAGCTCCTTGTGTGGCTGATGCTGTGGAATGATGTAGTTCAGAGAAAGCATTTTTTTT	1583		
DB	82086	TCAGCTCCTTGTGTGGCTGATGCTGTGGAATGATGTAGTTCAGAGAAAGCATTTTTTTT	82027		
QY	1584	TTCTTTGAACCTTAAAGTCTTATTTATTAAGCAGACAGATTTCCACATTTTTTATCAT	1643		
DB	82026	TTCTTTGAACCTTAAAGTCTTATTTATTAAGCAGACAGATTTCCACATTTTTTATCAT	81967		
QY	1644	GAGGATCTTCTTGGTGGTAAATACGAGATGACAGATCCCTTTAAAGAAGTTTATG	1703		
DB	81966	GAGGATCTTCTTGGTGGTAAATACGAGATGACAGATCCCTTTAAAGAAGTTTATG	81908		
QY	1704	TCCCTGACTCTGGCTAAATATCTAATTTCCAGATGCTTTTGTAGATGACTGAAGTAT	1763		
DB	81907	TCCCTGACTCTGGCTAAATATCTAATTTCCAGATGCTTTTGTAGATGACTGAAGTAT	81848		
QY	1764	TGTGAGCCACATATTGGGAGTCTAGATTTGAGTGAATGGCAGGAAAGGCCATCTCCAT	1823		
DB	81847	TGTGAGCCACATATTGGGAGTCTAGATTTGAGTGAATGGCAGGAAAGGCCATCTCCAT	81788		
QY	1824	TGAGATGATTAAGTGAACCAAACTAGTCTCGGAATTTACAGAGAGAGGGAATTCAGA	1883		
DB	81787	TGAGATGATTAAGTGAACCAAACTAGTCTCGGAATTTACAGAGAGAGGGAATTCAGA	81728		
QY	1884	CTGAGAGAGCTGTGACATAGGACTTGAAGACCAAAAGACTTTGAAATTTCCAGAGCTGCTCA	1943		
DB	81727	CTGAGAGAGCTGTGACATAGGACTTGAAGACCAAAAGACTTTGAAATTTCCAGAGCTGCTCA	81668		
QY	1944	TGTGTGAGTTATTTATCATGCTGTCTTTCTATTGAGTTACAATCTATATTTTATTCAA	2003		
DB	81667	TGTGTGAGTTATTTATCATGCTGTCTTTCTATTGAGTTACAATCTATATTTTATTCAA	81608		
QY	2004	GTTTAAATTAAGAAAAAATTTACAGA 2030			
DB	81607	GTTTAAATTAAGAAAAAATTTACAGA 81581			

```

DEFINITION Homo sapiens chromosome 1 clone RP11-204L3 map 1, *** SEQUENCING IN
ACCESSION AC025002
VERSION AC025002.3 GI:8516086
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Mammalia: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
AUTHORS Eukaryota: Eutheria: Primates: Catarrhini: Homnidae: Homo.
TITLE 1 (bases 1 to 200484)
JOURNAL Birren,B., Linton,L., Nusbaum,C. and Lander,E.
REFERENCE Homo sapiens chromosome 1, clone RP11-204L3
AUTHORS Unpublished
2 (bases 1 to 200484)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckhagalter,B., Brown,A., Burkett,G.,
Campiolo,A., Castle,A., Choepel,K., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Coyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Menes,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teasdale,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (03-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 14, 2000 this sequence version replaced gi:7770499.
All repeats were identified using RepeatMasker:
Smith,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7675
Center clone name: 204_L_3
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 81 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1048: contig of 1048 bp in length
* 1049 1148: gap of 100 bp
* 1149 2470: contig of 1322 bp in length
* 2471 2570: gap of 100 bp
* 2571 3643: contig of 1073 bp in length
* 3644 3743: gap of 100 bp
* 3744 4980: contig of 1237 bp in length
* 4981 5080: gap of 100 bp
* 5081 6352: contig of 1272 bp in length
* 6353 6452: gap of 100 bp
* 6453 7549: contig of 1097 bp in length
* 7550 7649: gap of 100 bp
* 7650 8977: contig of 1328 bp in length
* 8978 9077: gap of 100 bp
*
* 9078 10103: contig of 1026 bp in length
* 10104 10203: gap of 100 bp
* 10204 11362: contig of 1159 bp in length
* 11363 11462: gap of 100 bp
* 11463 12821: contig of 1359 bp in length
* 12822 12921: gap of 100 bp
* 12922 14007: contig of 1086 bp in length
* 14008 14107: gap of 100 bp
* 14108 15283: contig of 1176 bp in length
* 15284 15383: gap of 100 bp
* 15384 16399: contig of 1016 bp in length
* 16400 16499: gap of 100 bp
* 16500 17804: contig of 1305 bp in length
* 17805 17904: gap of 100 bp
* 17905 19176: contig of 1272 bp in length
* 19177 19276: gap of 100 bp
* 19277 20384: contig of 1108 bp in length
* 20385 20484: gap of 100 bp
* 20485 21689: contig of 1205 bp in length
* 21690 21789: gap of 100 bp
* 21790 23191: contig of 1402 bp in length
* 23192 23291: gap of 100 bp
* 23292 24513: contig of 1222 bp in length
* 24514 24613: gap of 100 bp
* 24614 26135: contig of 1522 bp in length
* 26136 26235: gap of 100 bp
* 26236 27373: contig of 1138 bp in length
* 27374 27473: gap of 100 bp
* 27474 28723: contig of 1250 bp in length
* 28724 28823: gap of 100 bp
* 28824 30217: contig of 1394 bp in length
* 30218 30317: gap of 100 bp
* 30318 32056: contig of 1739 bp in length
* 32057 32156: gap of 100 bp
* 32157 33818: contig of 1662 bp in length
* 33819 33918: gap of 100 bp
* 33919 35156: contig of 1238 bp in length
* 35157 35256: gap of 100 bp
* 35257 36712: contig of 1456 bp in length
* 36713 36812: gap of 100 bp
* 36813 38379: contig of 1567 bp in length
* 38380 38479: gap of 100 bp
* 38480 40190: contig of 1711 bp in length
* 40191 40290: gap of 100 bp
* 40291 41760: contig of 1490 bp in length
* 41761 41880: gap of 100 bp
* 41881 44053: contig of 2173 bp in length
* 44054 44153: gap of 100 bp
* 44154 45474: contig of 1321 bp in length
* 45475 45574: gap of 100 bp
* 45575 46821: contig of 1247 bp in length
* 46822 46921: gap of 100 bp
* 46922 47302: contig of 381 bp in length
* 47303 47402: gap of 100 bp
* 47403 49323: contig of 1921 bp in length
* 49324 49423: gap of 100 bp
* 49424 51140: contig of 1717 bp in length
* 51141 51240: gap of 100 bp
* 51241 53099: contig of 1859 bp in length
* 53100 53199: gap of 100 bp
* 53200 54631: contig of 1432 bp in length
* 54632 54731: gap of 100 bp
* 54732 55751: contig of 1020 bp in length
* 55752 55851: gap of 100 bp
* 55852 57470: contig of 1619 bp in length
* 57471 57570: gap of 100 bp
* 57571 60182: contig of 2612 bp in length
* 60183 60282: gap of 100 bp
* 60283 63065: contig of 2783 bp in length
* 63066 63165: gap of 100 bp
* 63166 64443: contig of 1278 bp in length
* 64444 64543: gap of 100 bp
* 64544 66043: contig of 1500 bp in length

```



```

* 66044 66143: gap of 100 bp
* 66144 68623: contig of 2480 bp in length
* 68624 68723: gap of 100 bp
* 68724 70122: contig of 1399 bp in length
* 70123 70222: gap of 100 bp
* 70223 72407: contig of 2185 bp in length
* 72408 72507: gap of 100 bp
* 72508 74635: contig of 2128 bp in length
* 74636 74735: gap of 100 bp
* 74736 77554: contig of 2819 bp in length
* 77555 77654: gap of 100 bp
* 77655 80277: contig of 2623 bp in length
* 80278 80377: gap of 100 bp
* 80378 81798: contig of 1421 bp in length
* 81799 81898: gap of 100 bp
* 81899 84126: contig of 2228 bp in length
* 84127 84226: gap of 100 bp
* 84227 86794: contig of 2568 bp in length
* 86795 86894: gap of 100 bp
* 86895 90311: contig of 3417 bp in length
* 90312 90411: gap of 100 bp
* 90412 92255: contig of 1844 bp in length
* 92256 92355: gap of 100 bp
* 92356 94570: contig of 2215 bp in length
* 94571 94670: gap of 100 bp
* 94671 97163: contig of 2493 bp in length
* 97164 97263: gap of 100 bp
* 97264 99724: contig of 2461 bp in length
* 99725 99824: gap of 100 bp
* 99825 103151: contig of 3327 bp in length
* 103152 103251: gap of 100 bp
* 103252 106370: contig of 3119 bp in length
* 106371 106470: gap of 100 bp
* 106471 108888: contig of 2418 bp in length
* 108889 108988: gap of 100 bp
* 108989 111514: contig of 2526 bp in length
* 111515 111614: gap of 100 bp
* 111615 114360: contig of 2746 bp in length
* 114361 114460: gap of 100 bp
* 114461 118269: contig of 3809 bp in length
* 118270 118369: gap of 100 bp
* 118370 122359: contig of 3990 bp in length
* 122360 122459: gap of 100 bp
* 122460 125669: contig of 3210 bp in length
* 125670 125769: gap of 100 bp
* 125770 129306: contig of 3537 bp in length
* 129307 129406: gap of 100 bp
* 129407 132984: contig of 3578 bp in length
* 132985 133084: gap of 100 bp
* 133085 137360: contig of 4276 bp in length
* 137361 137460: gap of 100 bp
* 137461 140476: contig of 3016 bp in length
* 140477 140576: gap of 100 bp

```

Query Match 29.1% Score 593; DB 2; Length 200484;
 Best Local Similarity 99.5%; Pred. No. 5.7e-313;
 Matches 863; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

QY 1164 CAGCCACATAAATCCAAATCCACTTGAACCACTGAAGAAAAGAGAATTGGACT 1223
    |||||
DB 94028 CAGCGACACTAAATCCAAATCCCACTTGAACCACTGAAGAAAAGAGAATTGGACT 93969
    |||||
QY 1224 CAGATGCTGTCAACACCATAGAAAGTCAAGTGTTCCTCCCACTTAACAAAGAACCTTA 1283
    |||||
DB 93968 CAGATGCTGTCAACACCATAGAAAGTCAAGTGTTCCTCCCACTTAACAAAGAACCTTA 93909
    |||||
QY 1284 AAGGGAAGGATCCCGAGGCGAATGATCAAGAGCTGAATGACGCTTTAAGACACAT 1343
    |||||
DB 93908 AAGAGAAAGGATCCCGAGGCGAATGATCAAGAGCTGAATGACGCTTTAAGACACAT 93849
    |||||
QY 1344 GTCCAGTGGCTTTGGAAGGTGTGGGGATCCAGTTCAAGAGGAGGAGGTATGTTTGC 1403
    |||||
DB 93848 GTCCAGTGGCTTTGGAAGGTGTGGGGATCCAGTTCAAGAGGAGGAGGTATGTTTGC 93789
    |||||

```

```

QY 1404 TCCGAGTCTGGGCAAGAGAGTGTCTATGCGGAATTCCTGTCATAGCAGAAAAGCTCCACC 1463
    |||||
DB 93788 TCCGAGTCTGGGCAAGAGAGTGTCTATGCGGAATTCCTGTCATAGCAGAAAAGCTCCACC 93729
    |||||
QY 1464 ATTTTCTTTTGTATGTTGTTTAAAGTCTACGCTTCTCTATATAGAAACAGCAGGCTCTTG 1523
    |||||
DB 93728 ATTTTCTTTTGTATGTTGTTTAAAGTCTACGCTTCTCTATATAGAAACAGCAGGCTCTTG 93669
    |||||
QY 1524 TCAGTCTCTGTGGGCGATGATGTCTGGAATATGATAGTATGATCAGAAAAGCTTTTCTT 1583
    |||||
DB 93668 TCAGTCTCTGTGGGCGATGATGTCTGGAATATGATAGTATGATCAGAAAAGCTTTTCTT 93609
    |||||
QY 1584 TTCTTTGAACCTTAAGGTTCTATTATTAAAGCAGACACAGATTCACATTTTATACAT 1643
    |||||
DB 93608 TTCTTTGAACCTTAAGGTTCTATTATTAAAGCAGACACAGATTCACATTTTATACAT 93549
    |||||
QY 1644 GAGATCTTCTTTGTTGTTGTAATACAGAGATTACTGTCATCCCTTTAAAGAAAGTTTATG 1703
    |||||
DB 93548 GAGATCTTCTTTGTTGTTGTAATACAGAGATTACTGTCATCCCTTT-AAAGAAAGTTTATG 93490
    |||||
QY 1704 TCCGAGCTCTGGCTAAATATTCTAATTTCCAGATGCTTTGTAGATCAGAGTAT 1763
    |||||
DB 93489 TCCGAGCTCTGGCTAAATATTCTAATTTCCAGATGCTTTGTAGATCAGAGTAT 93430
    |||||
QY 1764 TGTGAGCCACATATTGGAGGTTCTAGATTGAGTGAATGGCAGAAAGGCCATCTCAT 1823
    |||||
DB 93429 TGTGAGCCACATATTGGAGGTTCTAGATTGAGTGAATGGCAGAAAGGCCATCTCAT 93370
    |||||
QY 1824 TGAGATGATTAAAGTAACCAAACTAGTCTCGGAATTTCTACAGACAGAGGGAATCAGA 1883
    |||||
DB 93369 TGAGATGATTAAAGTAACCAAACTAGTCTCGGAATTTCTACAGAGAGGGAATCAGA 93310
    |||||
QY 1884 CTGAGAGAGCTGTGACATAGGACTGGAAGCAAGACACTTAAATTTGCGAGCTGCTCA 1943
    |||||
DB 93309 CTGAGAGAGCTGTGACATAGGACTGGAAGCAAGACACTTAAATTTGCGAGCTGCTCA 93250
    |||||
QY 1944 TGTGAGATTATTATACAGCTGCTGCTTTTATTGATGATTCATATATTTTATTGAA 2003
    |||||
DB 93249 TGTGAGATTATTATACAGCTGCTGCTTTTATTGATGATTCATATATTTTATTGAA 93190
    |||||
QY 2004 GTTTAAATTAAGAAAAATTTACAGA 2030
    |||||
DB 93189 GTTTAAATTAAGAAAAATTTACAGA 93163
    |||||
RESULT 10
AL354888 199863 bp DNA linear HTG 13-JUN-2001
LOCUS AL354888/C
DEFINITION Homo sapiens chromosome 1 clone RP11-473P22, v* SEQUENCING IN
VERSION AL354888.6 GI:9863692
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Plumb/B.
REFERENCE 1 (bases 1 to 199863)
AUTHORS Direct Submission
TITLE Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL CB10 15A, UK. E-mail enquiries: humquery@sanger.ac.uk
On Aug 21, 2000 this sequence version replaced gi:9213379.
COMMENT
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
Project Information
Center project name: ba473p22
Summary Statistics

```



```
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 10875; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 189476 bases at least Q40
Consensus quality: 193311 bases at least Q30
Consensus quality: 195510 bases at least Q20
Insert size: 198363; sum-of-contigs
Insert size: 201159; 6.2% error; agarose-ff
Quality coverage: 3.87x in Q20 bases; sum-of-contigs Quality
Coverage: 3.98x in Q20 bases; agarose-ff
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
2279 2378: contig of 2278 bp in length
2379 22743: contig of 20365 bp in length
22744 22843: gap of 100 bp
22844 45183: contig of 22340 bp in length
45184 45283: gap of 100 bp
45284 49581: contig of 4298 bp in length
49582 49681: gap of 100 bp
49682 58258: contig of 8577 bp in length
58259 58358: gap of 100 bp
58359 64082: contig of 5724 bp in length
64083 64182: gap of 100 bp
64183 72764: contig of 8582 bp in length
72765 72864: gap of 100 bp
72865 86552: contig of 13688 bp in length
86553 93481: contig of 6829 bp in length
93482 93581: gap of 100 bp
93582 101614: contig of 8033 bp in length
101615 101714: gap of 100 bp
101715 107836: contig of 6122 bp in length
107837 107936: gap of 100 bp
107937 144546: contig of 36610 bp in length
144547 144646: gap of 100 bp
144647 155892: contig of 11246 bp in length
155893 155992: gap of 100 bp
155993 167875: contig of 11883 bp in length
167876 167975: gap of 100 bp
167976 183411: contig of 15436 bp in length
183412 183511: gap of 100 bp
183512 199863: contig of 16352 bp in length.
-----
FEATURES
source
1..199863
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-473p22"
/clone_1lb="RP11-11.2"
1..2278
/feature="assembly-fragment:00765"
fragment_chain:1
clone_end:17
vector_side:left
2379..22743
/feature="assembly-fragment:01823"
fragment_chain:1
22844..45183
/feature="assembly-fragment:01809"
fragment_chain:2
45284..49581
/feature="assembly-fragment:01177"
fragment_chain:2
49682..58258
/feature="assembly-fragment:00999"
```

```
misc_feature      fragment_chain:2"
58359..64082
/feature="assembly-fragment:01187"
fragment_chain:2"
64183..72764
/feature="assembly-fragment:00292"
fragment_chain:2"
72865..86552
/feature="assembly-fragment:00339"
fragment_chain:3"
86553..93481
/feature="assembly-fragment:01156"
fragment_chain:3"
93482..101614
/feature="assembly-fragment:01960"
fragment_chain:3"
101715..107836
/feature="assembly-fragment:00867"
fragment_chain:4"
107937..144546
/feature="assembly-fragment:01896"
fragment_chain:4"
144647..155892
/feature="assembly-fragment:00307"
155993..167875
/feature="assembly-fragment:00528"
167976..183411
/feature="assembly-fragment:01170"
183512..199863
/feature="assembly-fragment:01894"
45915 c 45981 g 53154 t 1531 others
BASE COUNT      53282 a
ORIGIN
Query Match      23.2%; Score 473; DB 2; Length 199863;
Best Local Similarity 99.4%; Pred. No. 3.4e-247;
Matches 863; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
QY 1164 CACCGACATTAATATCAATCCACTTGAACCACTGAAGAAGAAGAGATTGGACT 1223
Db 26141 CACCGACATTAATATCAATCCACTTGAACCACTGAAGAAGAAGAAGATTGGACT 26082
QY 1224 CAGATCTGTCAACACCATAGAAAGTCAGAGTGTTCGCCAGACTATACCAAGAACCTA 1283
Db 26081 CAGATCTGTCAACACCATAGAAAGTCAGAGTGTTCGCCAGACTATACCAAGAACCTA 26022
QY 1284 AAGGGAAGGAGATCCCGACAGGCAATATGATCAAGAGCTGAAATCCAGCTTTAAGACACT 1343
Db 26021 AAGGGAAGGAGATCCCGACAGGCAATATGATCAAGAGCTGAAATCCAGCTTTAAGACACT 25962
QY 1344 GTCCAGTGGCTTTGGAAAGGTGTGGGATCCAGTTCAAGAGAGAGGATATGTTGTC 1403
Db 25961 GTCCAGTGGCTTTGGAAAGGTGTGGGATCCAGTTCAAGAGAGAGGATATGTTGTC 25902
QY 1404 TCCCACTGTGGCAAGAGATGCTATGCGAATTCCTGTGATAGCAGAAAAGCTCCACC 1463
Db 25901 TCCCACTGTGGCAAGAGATGCTATGCGAATTCCTGTGATAGCAGAAAAGCTCCACC 25842
QY 1464 ATTTCTTTTGAATGTGTTTAAAGTCTACGTTCTCTATATAGAAACAGCAGGCTTG 1523
Db 25841 ATTTCTTTTGAATGTGTTTAAAGTCTACGTTCTCTATATAGAAACAGCAGGCTTG 25782
QY 1524 TCAGCTCCTGTGTGGCTGTGTGCTGGAAGATGATGTTAGAGAAAGCATTTTCTT 1583
Db 25781 TCAGCTCCTGTGTGGCTGTGTGCTGGAAGATGATGTTAGAGAAAGCATTTTCTT 25722
QY 1584 TTCTTTGAACCTTAAGGTTCTATTATTAAGCAGACAGATTCACA-TTTTATPACA 1642
Db 25721 TTCTTTGAACCTTAAGGTTCTATTATTAAGCAGACAGATTCACA-TTTTATPACA 25662
QY 1643 TGAGGATCTTCTTGTGTGCTGAATACAGAGATTGATGCTATCCCTTTAAAGAAATTAT 1702
Db 25661 TGAGGATCTTCTTGTGTGCTGAATACAGAGATTGATGCTATCCCTT-AAAGAGTTTAT 25603
```

QY 1703 GTCCCTGACTGTGCTAAATTAATCTAATTTCCAGATGCTTTTGTAGATGATGAAGTAT 1762
 |||||||
 DB 25602 GTCCCTGACTGTGCTAAATTAATCTAATTTCCAGATGCTTTTGTAGATGATGAAGTAT 25543
 |||||||
 QY 1763 TTGTGAGCCACATATTGGAGCTTCTAGATTGTAGTGAATGGCGAAGGGCCATCTCCA 1822
 |||||||
 DB 25542 TTGTGAGCCACATATTGGAGCTTCTAGATTGTAGTGAATGGCGAAGGGCCATCTCCA 25483
 |||||||
 QY 1823 TTGAGATGATTAGTAGAACCACAACTAGTCTCGGAATTTACAGAGAGAGGGAATCAG 1882
 |||||||
 DB 25482 TTGAGATGATTAGTAGAACCACAACTAGTCTCGGAATTTACAGAGAGAGGGAATCAG 25423
 |||||||
 QY 1883 ACTGAGGAACCTGTGACATATGAGACTTGAAGACCAAGACTTTGAAATTTGCGAGCTGCTC 1942
 |||||||
 DB 25422 ACTGAGGAACCTGTGACATATGAGACTTGAAGACCAAGACTTTGAAATTTGCGAGCTGCTC 25363
 |||||||
 QY 1943 ATGTGTGAGTATTATTCACATGCTGCTCTTCTTATTTAGTACAAATCTATATTTTATTTGA 2002
 |||||||
 DB 25362 ATGTGTGAGTATTATTCACATGCTGCTCTTCTTATTTAGTACAAATCTATATTTTATTTGA 25303
 |||||||
 QY 2003 AGTTTAAATTAAGAAAAAATTTACAGA 2030
 |||||||
 DB 25302 AGTTTAAATTAAGAAAAAATTTACAGA 25275
 |||||||

RESULT 11
 AX336538/c
 LOCUS AX336538 355 bp DNA linear PAT 09-JAN-2002
 DEFINITION Sequence 7047 from Patent WO0194629.
 AX336538
 VERSION AX336538.1 GI:18127257
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 AUTHORS 1
 Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
 Horrigan, S., Sopper, D. R. and Weaver, Z.
 TITLE Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL Patent: WO 0194629-A 7047 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 source location/Qualifiers
 1..355
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 112 a 74 c 54 g 115 t
 ORIGIN

Query Match 11.8%; Score 241; DB 6; Length 355;
 Best Local Similarity 99.4%; Pred. No. 6; 8e-120;
 Matches 341; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1690 AAAGAACTTTATGTCCCTGACTGTGCTAAATTAATCTAATTTCCAGATGCTTTTGTAG 1749
 |||||||
 DB 343 AAAGAACTTTATGTCCCTGACTGTGCTAAATTAATCTAATTTCCAGATGCTTTTGTAG 284
 |||||||
 QY 1750 ATGACGTAAGTATTGTGAGCCACATATTGGAGCTTCTAGATTGTAGTGAATGGCAGAA 1809
 |||||||
 DB 283 ATGACGTAAGTATTGTGAGCCACATATTGGAGCTTCTAGATTGTAGTGAATGGCAGAA 224
 |||||||
 QY 1810 AGGCCCATCTTCATTTGAGATGATTAAGTGAACCAACTAGTTCGCGAATTTCTACAGAA 1869
 |||||||
 DB 223 AGGCCCATCTTCATTTGAGATGATTAAGTGAACCAACTAGTTCGCGAATTTCTACAGAA 164
 |||||||
 QY 1870 AGAGGAATCTAGACTAGAGAACTGTGACATAGACTTGAAGCAACCAAGACTTTGAAT 1929
 |||||||
 DB 163 AGAGGAATCTAGACTAGAGAACTGTGACATAGACTTGAAGCAACCAAGACTTTGAAT 104
 |||||||
 QY 1930 TTGCGAGCTGCTCATGTGTGAGTATTATATCACTGCTCTCTTTCTATTGAGTTACAAATCT 1989
 |||||||
 DB 103 TTGCGAGCTGCTCATGTGTGAGTATTATATCACTGCTCTCTTTCTATTGAGTTACAAATCT 44
 |||||||

QY 1990 ATATTTTATTGAGTTTAAATTAAGAAAAAATTTACAGAAA 2032
 |||||||
 DB 43 ATATTTTATTGAGTTTAAATTAAGAAAAAATTTACAGAAA 1

RESULT 12
 G24438/c
 LOCUS G24438 252 bp DNA linear STS 31-MAY-1996
 DEFINITION human STS WI-12773, sequence tagged site.
 G24438
 ACCESSION G24438.1 GI:1344764
 VERSION
 KEYWORDS STS; STS sequence; primer; sequence tagged site.
 SOURCE Homo sapiens STS derived from sequences in dbEST and the Unigene
 collection.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 AUTHORS 1 (bases 1 to 252)
 TITLE Whitehead Institute/MIT Center for Genome Research; Physically
 Mapped STS
 JOURNAL Unpublished (1995)
 COMMENT

Contact: Thomas Hudson
 Whitehead Institute/MIT Center for Genome Research
 Whitehead Institute for Biomedical Research
 9 Cambridge Center, Cambridge MA 02142 USA
 Tel: 617 252 1900
 Fax: 617 252 1902
 Email: thudson@genome.wi.mit.edu

Primer A: TAATCTCATAGAAAGACGACGTGA
 Primer B: AGTTCTAGATTGAGTGAATGCA
 STS size: 201
 PCR Profile:

Presoak:
 Denaturation:
 Annealing: 56 degrees C
 Polymerization:
 PCR Cycles: 35
 Thermal Cycler:
 Protocol:
 Template: 10 ng
 Primer: each 5 pm
 dNTPs: each 4 mM
 Taq Polymerase: 0.025 units/ul
 Total Vol: 20 ul

Buffer:
 MgCl2: 1.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 9.3

FEATURES
 source Derived from dbEST (genbank accession Z40724).
 location/Qualifiers
 1..252

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="122.8 CR from top of Chr1 linkage group"
 STS
 primer_bind 48..72
 primer_bind complement(225..248)
 BASE COUNT 74 a 58 c 34 g 85 t 1 others
 ORIGIN

Query Match 7.3%; Score 150; DB 11; Length 252;
 Best Local Similarity 99.2%; Pred. No. 5e-70;
 Matches 250; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1778 TGGGAGTCTAGATTGAGTGAATGCGAAGAGCCCATCTCCATTGAGATTAATTAAGT 1837
 |||||||

DB	252	TGGGAGTTCAGATTGAGTCAAGTATGGCAGAAAGGCCATCTCCATTGAGATGATTAAGT	193
OY	1838	GAACCAACTAGTTCGGAATTCACAGAGAAGGAGGATCAGACTGAGAGCTGTCG	1897
DB	192	GAACCAACTAGTTCGGAATTCACAGAGAAGGAGGATCAGACTGAGAGCTGTCG	133
OY	1898	ACATGAGACTTGAGACCAAGACTTTGAAATTTGCGAGCTGCATGCTGATATTA	1957
DB	132	ACATGAGACTTGAGACCAAGACTTTGAAATTTGCGAGCTGCATGCTGATATTA	73
OY	1958	TCACGCGTCTCTTCTATGAGTTCACAAATCATATTTTATGAGATTAAATAAGAA	2017
DB	72	TCACGCGTCTCTTCTATGAGTTCACAAATCATATTTTATGAGATTAAATAAGAA	13
OY	2018	AAATTTACAG 2029	
DB	12	AAATTTACAG 1	
RESULT 13			
AL354888		199863 bp	DNA
LOCUS			linear
DEFINITION			HTG_13-JUN-2001
ACCESSION	AL354888		
VERSION	AL354888.6		
KEYWORDS	HTG; HTGS; PHASE1; HTGS_CANCELLED.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 199863)		
JOURNAL	Plumb, B.		
COMMENT	Direct Submission		
	Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,		
	CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk		
	requests: clonerequests@sanger.ac.uk		
	On Aug 21, 2000 this sequence version replaced gi:9213379.		
	Genome Center		
	Center: Sanger Centre		
	Center code: SC		
	Web site: http://www.sanger.ac.uk		
	Contact: humquerry@sanger.ac.uk		
	Project Information		
	Center project name: BA473P22		
	Summary Statistics		
	Assembly program: XGAP4; version 4.5		
	Sequencing vector: plasmid; L08752; 100% of reads		
	Chemistry: Dye-terminator Big Dye; 100% of reads		
	Consensus quality: 189476 bases at least Q40		
	Consensus quality: 193311 bases at least Q30		
	Consensus quality: 195510 bases at least Q20		
	Insert size: 198363; sum-of-configs		
	Insert size: 201159; 6.2% error; agarose-fp		
	Quality coverage: 3.87x in Q20 bases; sum-of-configs Quality		
	coverage: 3.98x in Q20 bases; agarose-fp		
	NOTE: This is a 'working draft' sequence. It currently		
	consists of 16 contigs. The true order of the pieces		
	is not known and their order in this sequence record is		
	arbitrary. Gaps between the contigs are represented as		
	runs of N, but the exact sizes of the gaps are unknown.		
	This record will be updated with the finished sequence		
	as soon as it is available and the accession number will		
	be preerved.		
	1 2278: contig of 2278 bp in length		
	2279 2378: gap of 100 bp		
	2379 22743: contig of 20365 bp in length		
	22744 22843: gap of 100 bp		
	22844 45183: contig of 22340 bp in length		
	45184 45283: gap of 100 bp		
	45284 49581: contig of 4298 bp in length		
	49582 49681: gap of 100 bp		

*	49682	58258:	contig of 8577 bp in length
*	58259	58358:	gap of 100 bp
*	58359	64082:	contig of 5724 bp in length
*	64083	64182:	gap of 100 bp
*	64183	72764:	contig of 8582 bp in length
*	72765	72864:	gap of 100 bp
*	72865	86552:	contig of 13688 bp in length
*	86553	86652:	gap of 100 bp
*	86653	93481:	contig of 6829 bp in length
*	93482	93581:	gap of 100 bp
*	93582	101614:	contig of 8033 bp in length
*	101615	101714:	gap of 100 bp
*	101715	107836:	contig of 6122 bp in length
*	107837	107936:	gap of 100 bp
*	107937	144547:	contig of 11246 bp in length
*	144547	144646:	gap of 100 bp
*	144647	155892:	contig of 11246 bp in length
*	155893	155992:	gap of 100 bp
*	155993	167875:	contig of 11883 bp in length
*	167876	167975:	gap of 100 bp
*	167976	183411:	contig of 15436 bp in length
*	183412	183511:	gap of 100 bp
*	183512	199863:	contig of 16352 bp in length
FEATURES	Location/Qualifiers		
source	1..199863		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/clone="RP11-473p22"		
	/clone_1lb="RPCT-11.2"		
	1..2278		
misc_feature	/note="assembly_fragment:00765 fragment_chain:1 clone_end:77 vector_side:left"		
	2379..22743		
misc_feature	/note="assembly_fragment:01823 fragment_chain:1"		
	22844..45183		
misc_feature	/note="assembly_fragment:01809 fragment_chain:2"		
	45284..49581		
misc_feature	/note="assembly_fragment:01177 fragment_chain:2"		
	49682..58258		
misc_feature	/note="assembly_fragment:00999 fragment_chain:2"		
	58359..64082		
misc_feature	/note="assembly_fragment:01187 fragment_chain:2"		
	64183..72764		
misc_feature	/note="assembly_fragment:00292 fragment_chain:2"		
	72865..86552		
misc_feature	/note="assembly_fragment:00339 fragment_chain:3"		
	86653..93481		
misc_feature	/note="assembly_fragment:01156 fragment_chain:3"		
	93482..101614		
misc_feature	/note="assembly_fragment:01960 fragment_chain:3"		
	101715..107836		
misc_feature	/note="assembly_fragment:00867 fragment_chain:4"		
	107937..144546		
misc_feature	/note="assembly_fragment:01896 fragment_chain:4"		
	144647..155892		
misc_feature	/note="assembly_fragment:00307"		
	155993..167875		
misc_feature	/note="assembly_fragment:00528"		
	167976..183411		

```

/Note="assembly_fragment:01170"
misc.feature      183512..199863
                  /Note="assembly_fragment:01894"
BASE COUNT       53282 a 45915 c 45981 g 53154 t 1531 others
ORIGIN
Query Match      7 0%: Score 143; DB 2; Length 199863;
Best Local Similarity 100.0%; Pred. No. 2,1e+66;
Matches 143: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 44 GTGGGAGTGGGCTCAGGGGCTGCACAGGACCTACCTCTTGTAGTATCTCGGGCC 103
      |||
DB 8752 GTGGGAGTGGGCTCAGGGGCTGCACAGGACCTACCTCTTGTAGTATCTCGGGCC 8811
      |||

OY 104 ACGGGACACCGCAATCTCAGCTGCTTCAGCTAGGCCAGCGGCTCGGCGATC 163
      |||
DB 8812 ACGGGACACCGCAATCTCAGCTGCTTCAGCTAGGCCAGCGGCTCGGCGATC 8871
      |||

OY 164 CTCAGCGCTGACCTCATGCAAGT 186
      |||
DB 8872 CTCAGCGCTGACCTCATGCAAGT 8894
      |||

RESULT 14
AC025068/c       79785 bp   DNA   linear   HTG 13-JUL-2000
LOCUS            AC025068/c
DEFINITION      Homo sapiens chromosome 1 clone RP11-21702 map 1, LOW-PASS SEQUENCE
ACCESSION       AC025068
VERSION         AC025068.2 GI:7329438
KEYWORDS        HTG; HTGS_PHASE0.
SOURCE          Homo sapiens.
ORGANISM        Homo sapiens
REFERENCE       1 (bases 1 to 79785)
AUTHORS         Bliren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE           Homo sapiens chromosome 1, clone RP11-21702
JOURNAL         Unpublished
REFERENCE       2 (bases 1 to 79785)
AUTHORS         Bliren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
JOURNAL         Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
REFERENCE       Boucslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
AUTHORS         Campoliano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
JOURNAL         Collamore,A., Cooke,P., DeRetliano,K., Dewar,K., Diaz,J.S.,
REFERENCE       Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
AUTHORS         Galagan,J., Gardyna,S., Glnde,S., Goyette,M., Graham,L.,
JOURNAL         Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
REFERENCE       Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
AUTHORS         Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
JOURNAL         McCarthy,M., McEwan,P., McGuff,A., McKernan,K., McPheters,R.,
REFERENCE       McElrath,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
AUTHORS         Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
JOURNAL         O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
REFERENCE       Pisanic,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
AUTHORS         Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
JOURNAL         Stange-Thomann,N., Stojanovic,N., Sudarmanian,A., Talamas,J.,
REFERENCE       Testafaye,S., Theodore,J., Thirell,A., Travers,M., Trigilio,J.,
AUTHORS         Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
JOURNAL         Young,G., Zainoun,J., Zimmer,A. and Zody,M.
COMMENT         Direct Submission
                  Submitted (04-MAR-2000) Whitehead Institute/MIT Center for Genome
                  Research, 320 Charles Street, Cambridge, MA 02141, USA
                  On Mar 26, 2000 this sequence version replaced g1:7158126.
                  All repeats were identified using RepeatMasker:
                  Smit, A.F.A. & Green, P. (1996-1997)
                  http://ftp.genome.washington.edu/RM/RepeatMasker.html
                  ----- Genome Center
                  Center: Whitehead Institute/ MIT Center for Genome Research
                  Center code: WIBR
                  Web site: http://www-seq.wi.mit.edu
                  Contact: sequence_submissions@genome.wi.mit.edu

```

```

----- Project Information
Center Project name: L7716
Center clone name: 217_0_2

```

NOTE: This record contains 98 individual sequencing reads that have not been assembled into contigs. Runs of N are used to separate the reads and the order in which they appear is completely arbitrary. Low-pass sequence sampling is useful for identifying clones that may be gene-rich and allow overlap relationships among clones to be deduced. However, it should not be assumed that this clone will be sequenced to completion. In the event that the record is updated, the accession number will be preserved.

```

1 740: contig of 740 bp in length
741 840: gap of 100 bp
841 1543: contig of 703 bp in length
1544 1643: gap of 100 bp
1644 2350: contig of 707 bp in length
2351 2450: gap of 100 bp
2451 3158: contig of 708 bp in length
3159 3258: gap of 100 bp
3259 3978: contig of 720 bp in length
3979 4078: gap of 100 bp
4079 4776: contig of 698 bp in length
4777 4876: gap of 100 bp
4877 5593: contig of 717 bp in length
5594 5693: gap of 100 bp
5694 6412: contig of 719 bp in length
6413 6512: gap of 100 bp
6513 7238: contig of 726 bp in length
7239 7338: gap of 100 bp
7339 8055: contig of 717 bp in length
8056 8155: gap of 100 bp
8156 8870: contig of 715 bp in length
8871 8970: gap of 100 bp
8971 9691: contig of 721 bp in length
9692 9791: gap of 100 bp
9792 10493: contig of 702 bp in length
10494 10593: gap of 100 bp
10594 11326: contig of 733 bp in length
11327 11426: gap of 100 bp
11427 12137: contig of 711 bp in length
12138 12237: gap of 100 bp
12238 12950: contig of 713 bp in length
12951 13050: gap of 100 bp
13051 13751: contig of 701 bp in length
13752 13851: gap of 100 bp
13852 14552: contig of 701 bp in length
14553 14652: gap of 100 bp
14653 15372: contig of 720 bp in length
15373 15472: gap of 100 bp
15473 16183: contig of 711 bp in length
16184 16283: gap of 100 bp
16284 17005: contig of 722 bp in length
17006 17105: gap of 100 bp
17106 17825: contig of 720 bp in length
17826 17925: gap of 100 bp
17926 18639: contig of 714 bp in length
18640 18739: gap of 100 bp
18740 19457: contig of 718 bp in length
19458 19557: gap of 100 bp
19558 20269: contig of 712 bp in length
20270 20369: gap of 100 bp
20370 21069: contig of 700 bp in length
21070 21169: gap of 100 bp
21170 21895: contig of 726 bp in length
21896 21995: gap of 100 bp
21996 22709: contig of 714 bp in length
22710 22809: gap of 100 bp
22810 23507: contig of 658 bp in length
23508 23607: gap of 100 bp

```

```

* 23608 24297: contig of 690 bp in length
* 24298 24397: gap of 100 bp
* 24398 25107: contig of 710 bp in length
* 25108 25207: gap of 100 bp
* 25208 25936: contig of 729 bp in length
* 25937 26036: gap of 100 bp
* 26037 26749: contig of 713 bp in length
* 26750 26849: gap of 100 bp
* 26850 27562: contig of 713 bp in length
* 27563 27662: gap of 100 bp
* 27663 28369: contig of 707 bp in length
* 28370 28469: gap of 100 bp
* 28470 29153: contig of 684 bp in length
* 29154 29253: gap of 100 bp
* 29254 29971: contig of 718 bp in length
* 29972 30071: gap of 100 bp
* 30072 30770: contig of 699 bp in length
* 30771 30870: gap of 100 bp
* 30871 31581: contig of 711 bp in length
* 31582 31681: gap of 100 bp
* 31682 32380: contig of 699 bp in length
* 32381 32480: gap of 100 bp
* 32481 33187: contig of 707 bp in length
* 33188 33287: gap of 100 bp
* 33288 34012: contig of 725 bp in length
* 34013 34112: gap of 100 bp
* 34113 34831: contig of 719 bp in length
* 34832 34931: gap of 100 bp
* 34932 35627: contig of 696 bp in length
* 35628 35727: gap of 100 bp
* 35728 36446: contig of 719 bp in length
* 36447 36546: gap of 100 bp
* 37246 37345: gap of 100 bp
* 37346 38077: contig of 732 bp in length
* 38078 38177: gap of 100 bp
* 38178 38879: contig of 702 bp in length
* 38880 38979: gap of 100 bp
* 38980 39688: contig of 709 bp in length
* 39689 39788: gap of 100 bp
* 39789 40500: contig of 712 bp in length
* 40501 40600: gap of 100 bp
* 40601 41317: contig of 717 bp in length
* 41318 41417: gap of 100 bp
* 41418 42116: contig of 699 bp in length
* 42117 42216: gap of 100 bp
* 42217 42923: contig of 707 bp in length
* 42924 43023: gap of 100 bp
* 43024 43739: contig of 716 bp in length
* 43740 43839: gap of 100 bp
* 43840 44570: contig of 731 bp in length
* 44571 44670: gap of 100 bp
* 44671 45381: contig of 711 bp in length
* 45382 45481: gap of 100 bp
* 45482 46191: contig of 710 bp in length
* 46192 46291: gap of 100 bp
* 46292 46986: contig of 695 bp in length
* 46987 47086: gap of 100 bp
* 47087 47789: contig of 703 bp in length
* 47790 47889: gap of 100 bp
* 47890 48590: contig of 701 bp in length
* 48591 48690: gap of 100 bp
* 48691 49414: contig of 724 bp in length
* 49415 49514: gap of 100 bp
* 49515 50220: contig of 706 bp in length
* 50221 50320: gap of 100 bp
* 50321 51010: contig of 690 bp in length
* 51011 51110: gap of 100 bp
* 51111 51823: contig of 713 bp in length
* 51824 51923: gap of 100 bp
* 51924 52648: contig of 725 bp in length
* 52649 52748: gap of 100 bp
* 52749 53466: contig of 718 bp in length

```

```

Query Match      5.9%; Score 120; DB 2; Length 79785;
Best Local Similarity 100.0%; Pred. No. 9,1e-54;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

* 53467 53566: gap of 100 bp
* 53567 54295: contig of 729 bp in length
* 54296 54395: gap of 100 bp
* 54396 55105: contig of 710 bp in length
* 55106 55205: gap of 100 bp
* 55206 55935: contig of 730 bp in length

```

```

OY 1047 AGGCCACAGCCCTACAGCCATCTCATTAAGATGGCATCAATGAGAGACAGAGA 1106
      |||||||
Db 2541 AGGCCACAGCCCTACAGCCATCTCATTAAGATGGCATCAATGAGAGACAGAGA 25582
      |||||||
OY 1107 GAAAGTTATCACCTGTGTGACCTCTGTGATGCAATCATCATTTGGGATCGGCAATGGGCGAG 1166
      |||||||
Db 25581 GAAAGTTATCACCTGTGTGACCTCTGTGATGCAATCATCATTTGGGATCGGCAATGGGCGAG 25522
      |||||||

RESULT 15
BC019812
LOCUS
DEFINITION Mus musculus, RIKEN CDNA 2310075G14 gene, clone MGC:30541
ACCESSION BC019812
VERSION BC019812.1 GI:18044185
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Mus musculus.
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 2045)
REFERENCE
  1 Strausberg, R.
  TITLE Direct Submission
  JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian
  Gene Collection (MGC), Cancer Genomics Office, National Cancer
  Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
  USA
REMARK
  NIH-MGC Project URL: http://mgc.nci.nih.gov
  Contact: MGC help desk
  Email: gcqabs-remail.nih.gov
  Tissue Procurement: Jeffrey E. Green, M.D.
  CDNA Library Preparation: Life Technologies, Inc.
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Baylor College of Medicine Human Genome
  Sequencing Center
  Center code: BCM-HGSC
  Web site: http://www.hgsc.bcm.tmc.edu/cdna/
  Contact: amgebcm.tmc.edu
  Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,
  Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
  Richards, S., Gibbs, R.A.

```

```

FEATURES
  source
    1..2045
      /organism="Mus musculus"
      /db_xref="taxon:10090"
      /map="FVB/N"
      /clone="MGC:30541 IMAGE:5042856"
      /tissue-type="Kidney, normal. 5 month old male mouse."
      /clone_lib="NCI-CGAP_K1d14"
      /lab_host="DH10B"
      /note="Vector: pCMV-SPORT6"
      406..1386
        /codon_start=1
        /product="RIKEN CDNA 2310075G14 gene"
        /protein_id="AAH19812.1"
        /db_xref="GI:18044186"
        /db_xref="LocusID:66966"

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 41 Row: 1 Column: 20.
 Location/Qualifiers
 1..2045

/translation="MGTGKVVDRKVELLEKEDGHELHKRLSQVDPENAAKLHPHDKRKV
ARSLQVFEETGISHSSEFLHROHAEEGGCPGGLRFPNPCIILWLHADQAVLDERLDKR
VDDMLAAGLLEELRGFRRRLNKLINSODYOHGIFOSIGFKEFHEYLITTEGKCTPE
TSNQLKKGTEALKQYTKRYARKQNRWVKNRRLSRPGPSVPPVYGLVSDVSKWEEVS
LEPALNIVOSFTIOGKPTAPVYKMAYNESBNKRSYHMCDCDRITITGDRMAHHLKSK
SHLQDKRRRLDLDVAVSATGSNSPDCDPERIEGESSGQHNOELKASY"
BASE COUNT 570 a 453 c 541 g 481 t
ORIGIN

Query Match 2.48; Score 48; DB 10; Length 2045;
Best Local Similarity 100.0%; Pred. No. 3.3e-14;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 365 GTGGAGGAGCAATTTACATGCAATCCTGCTGGAAGTTCTT 412
DB 337 GTGGAGGAGCAATTTACATGCAATCCTGCTGGAAGTTCTT 384

Search completed: April 21, 2003, 21:27:10
Job time : 6769 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 19:23:36 ; Search time 3289 seconds

(Without alignments)
10050.164 Million cell updates/sec

Title: US-09-513-151-3
Sequence: 1 CTCCCATAGATGCGCTCCG.....TTTACAGAAAAA 2041

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues.

Word size: 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST: *
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	573	28.1	725	9	AI133396 HA2003 Hu
2	569	27.9	743	14	BM975520 UT-CF-EN1
3	541	26.5	578	14	BM847727 K-EST0127
4	529	25.9	725	10	BE315223 BE315223
5	480	23.5	601	9	AI207688 HA3066 Hu
6	466	22.8	799	13	BI758117 BI758117

7	451	22.1	1088	14	BM800217
8	433	21.2	869	12	BG612651
9	422	20.7	639	10	AV759288
10	388	19.0	559	9	AI718450
11	387	19.0	624	13	B0003256
12	387	19.0	783	14	BI222788
13	385	18.9	457	9	AA121465
14	372	18.2	543	9	AI193562
15	371	18.2	562	12	BG495857
16	371	18.2	963	13	BI462398
17	370	18.1	490	9	AI127851
18	367	18.0	487	9	AI337829
19	367	18.0	724	14	BM721352
20	364	17.8	758	13	BI820461
21	361	17.7	469	13	BF376938
22	357	17.5	528	12	BF064275
23	352	17.2	430	9	AI089239
24	350	17.1	350	9	AA847885
25	350	17.1	465	9	AA767411
26	342	16.8	513	10	AA888867
27	342	16.8	650	9	AA104415
28	340	16.7	440	14	N26804
29	340	16.7	444	14	N26803
30	340	16.7	473	14	N29610
31	339	16.6	442	9	AA830847
32	338	16.6	413	10	AA193859
33	338	16.6	443	14	N41610
34	337	16.5	386	10	AA197736
35	337	16.5	446	9	AA121308
36	335	16.4	433	9	AA716699
37	329	16.1	500	9	AI192804
38	323	15.8	711	12	BG506209
39	321	15.7	643	10	AAW53634
40	320	15.7	554	9	AI749267
41	314	15.4	772	12	BG034660
42	309	15.1	836	12	BG575745
43	305	14.9	922	9	AL530575
44	301	14.7	434	9	AI347522
45	301	14.7	593	10	AA444470

ALIGNMENTS

RESULT 1
LOCUS AI133396/c 725 bp mRNA linear EST 11-NOV-1999
DEFINITION HA2003 Human fetal liver CDNA library Homo sapiens CDNA, mRNA
ACCESSION AI133396
VERSION AI133396.1 GI:6360712
KEYWORDS
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Yu, Y., Zhang, C., Luo, L., Ouyang, S., Li, W., Wu, J., Zhou, S., Liu, M. and He, F.
Expression profile analysis of a human fetal liver CDNA library
Unpublished (1998)
CONTACT Contact: Yongtao Yu
Department of Hematology
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100650, P.R.China
Tel: 0086-10-68159479
Fax: 0086-10-68214653
Email: yyl48@yahoo.com.

FEATURES

1. 725
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="Human fetal liver CDNA library"

```

/tissue_type="liver"
/dev_stage="fetal"
/lab_host="MC1061/p3"
/notes="Vector: pCDNA1"
BASE COUNT      220 a      177 c      124 g      204 t
ORIGIN

Query Match      28.1%; Score 573; DB 9; Length 725;
Best Local Similarity 99.7%; Pred. No. 2.8e-202;
Matches 693; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1296 CCCCAGGCGAGAAATGATCAAGAGCTGAATGACGCGTTTAAAGACATGTCAGTGCGCT 1355
      |||||||
DB 694 CCCCAGGCGAGAAATGATCAAGAGCTGAATGACGCGTTTAAAGACATGTCAGTGCGCT 635
      |||||||
OY 1356 TTGGAAGGCTGCGGGGATCCAGTTCAGAGGAGGAGGATGTTTGTCTCCAGTGGG 1415
      |||||||
DB 634 TTGGAAGGCTGCGGGGATCCAGTTCAGAGGAGGAGGATGTTTGTCTCCAGTGGG 575
      |||||||
OY 1416 CAAAGAGTCTGTCGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1475
      |||||||
DB 574 CAAAGAGTCTGTCGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 515
      |||||||
OY 1476 TGTGCTTTTAAAGTCTCAGCTCTCTATATATAGAAAACAGAGGCTCTGACCTCTGT 1535
      |||||||
DB 514 TGTGCTTTTAAAGTCTCAGCTCTCTATATATAGAAAACAGAGGCTCTGACCTCTGT 455
      |||||||
OY 1536 GTGCGTGAATGTCGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1595
      |||||||
DB 454 GTGCGTGAATGTCGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 395
      |||||||
OY 1596 TAAAGCTCTATTTTAAAGAGCAGACAGATTCACATTTTATACATGAGATCTTCT 1655
      |||||||
DB 394 TAAAGCTCTATTTTAAAGAGCAGACAGATTCACATTTTATACATGAGATCTTCT 335
      |||||||
OY 1656 TGTGCGTGAATGTCGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1715
      |||||||
DB 334 TGTGCGTGAATGTCGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 276
      |||||||
OY 1716 GCTAAATATCTAATTTTCCAGATGCTTTTGTAGATGATGATGATGATGATGATGAT 1775
      |||||||
DB 275 GCTAAATATCTAATTTTCCAGATGCTTTTGTAGATGATGATGATGATGATGATGAT 216
      |||||||
OY 1776 ATTGGAGTCTAGATTTTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1835
      |||||||
DB 215 ATTGGAGTCTAGATTTTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 156
      |||||||
OY 1836 GTGAACCAACTAGTCTCTGGAATTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAG 1895
      |||||||
DB 155 GTGAACCAACTAGTCTCTGGAATTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAG 96
      |||||||
OY 1896 TGACATAGAGCTTGAAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1955
      |||||||
DB 95 TGACATAGAGCTTGAAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 36
      |||||||
OY 1956 TATCACTGCTGCTCTTCTATTTAGATTACAATCT 1989
      |||||||
DB 35 TATCACTGCTGCTCTTCTATTTAGATTACAATCT 2
      |||||||

RESULT 2
BM975520/c      743 bp      mRNA      linear      EST 21-MAR-2002
LOCUS          UI-CF-EN1-acw-c-07-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
DEFINITION    UI-CF-EN1-acw-c-07-0-UI 3', mRNA sequence.
ACCESSION    BM975520
VERSION      BM975520.1 GI:19593111
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 743)

```

```

AUTHORS      Ronaldo, M.F., Lennon, G. and Soares, M.B.
TITLE        Normalization and subtraction: two approaches to facilitate gene
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      97044477
COMMENT      Contact: McCray, PB
               University of Iowa
               2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
               Tel: 319 356 4866
               Fax: 319 356 7171
               Email: paul-mccray@uiowa.edu
               Tissue Procurement: Dr. M. J. Welsh, University of Iowa
               cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
               DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
               Clone Distribution: Researchers may obtain clones from Research
               Genetics (www.resgen.com).
               The following repetitive elements were found in this cDNA
               sequence: 1-56, >AT-rich#low_complexity (matched complement)
               Seq primer: M13 FORWARD
               Polya=yes.

FEATURES
    source          location/Qualifiers
    1..743
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="UI-CF-EN1-acw-c-07-0-UI"
    /clone_11b="UI-CF-EN1"
    /tissue_type="Primary Lung Cystic Fibrosis Epithelial
    Cells"
    /dev_stage="Adult"
    /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
    /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
    modified polylinker; Site_1: EcoR I; Site_2: Not I;
    UI-CF-EN1 is a normalized cDNA library containing the
    following tissue(s): Primary Lung Cystic Fibrosis
    Epithelial Cells. The library was constructed according to
    Ronaldo, Lennon and Soares, Genome Research, 6:791-806,
    1996. First strand cDNA synthesis was primed with an
    oligo-dT primer containing a Not I site. Double stranded
    cDNA was ligated to an EcoR I adaptor, digested with Not
    I, and cloned directionally into pT73-Pac vector. The
    oligonucleotide used to prime the synthesis of
    first-strand cDNA contains a library tag sequence that is
    located between the Not I site and the (drr)18 tail. The
    sequence tag for this library is CTGCTCAGT.
    TAG_L1B=UI-CF-EN1
    TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
    6hr to LPS 24h
    TAG_SEQ=CTGCTCAGGT"

BASE COUNT      231 a      172 c      116 g      224 t
ORIGIN

Query Match      27.9%; Score 569; DB 14; Length 743;
Best Local Similarity 99.7%; Pred. No. 8.3e-201;
Matches 739; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1301 GGGCGAATGATCAAGAGCTGAATGACGCGTTTAAAGACATGTCAGTGCGCTTTGGA 1360
      |||||||
DB 743 GGGCGAATGATCAAGAGCTGAATGACGCGTTTAAAGACATGTCAGTGCGCTTTGGA 684
      |||||||
OY 1361 AAGCTGCTGCGGATCCAGTTCAGAGGAGGAGGATGTTTGTCTCCAGTCTGGGCAAG 1420
      |||||||
DB 683 AAGCTGCTGCGGATCCAGTTCAGAGGAGGAGGATGTTTGTCTCCAGTCTGGGCAAG 624
      |||||||
OY 1421 GAGTCTATGCGGAATTTCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1480
      |||||||
DB 623 GAGTCTATGCGGAATTTCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 564
      |||||||
OY 1481 TTTTAAAGTCTCAGCTTCTATATATAGAAAACAGAGAGCTTTGTCAGCTCTGTCGGC 1540
      |||||||
DB 563 TTTTAAAGTCTCAGCTTCTATATATAGAAAACAGAGAGCTTTGTCAGCTCTGTCGGC 504
      |||||||

```



```

OY 1541 TGATGTCMCGGAATGATGATGTCAGGAAGCATTTTCTTCTGAACCTTAAG 1600
    |||||||
DB 503 TGATGTCGTCGGAATGATGATGTCAGGAAGCATTTTCTTCTGAACCTTAAG 444
OY 1601 GTTCTATTATTAAGCAGCAGATTCACATTTTATACATGAGATCTTCTTGTGG 1660
    |||||||
DB 443 GTTCTATTATTAAGCAGCAGATTCACATTTTATACATGAGATCTTCTTGTGG 384
OY 1661 TGAATACAGGATTCAGTGCATCCCTTTAAAGAAGTTTATGTCCTGACCTGGCTAA 1720
    |||||||
DB 383 TGAATACAGGATTCAGTGCATCCCTTT -AAAGAAGTTTATGTCCTGACCTGGCTAA 325
OY 1721 AATTATCTAATTTCCAGATGCTTTGTGATGACGAGATTTGTGAGCCACATATTGG 1780
    |||||||
DB 324 AATTATCTAATTTCCAGATGCTTTGTGATGACGAGATTTGTGAGCCACATATTGG 265
OY 1781 GAGTCTGATTTGAGTAATGCGAGAAAGGCCATCTCCATTTGATGATTAAGTAA 1840
    |||||||
DB 264 GAGTCTGATTTGAGTAATGCGAGAAAGGCCATCTCCATTTGATGATTAAGTAA 205
OY 1841 CCAAACTAGTCTCGGAATTTCTACAGAGAGAGGGAATCAGACTGAGGAAGCTGTGACA 1900
    |||||||
DB 204 CCAAACTAGTCTCGGAATTTCTACAGAGAGAGGGAATCAGACTGAGGAAGCTGTGACA 145
OY 1901 TAGGACTGAGAGACCAAGACTTTGAATTTGCGAGCTGCTCATGAGTATTTATCA 1960
    |||||||
DB 144 TAGGACTGAGAGACCAAGACTTTGAATTTGCGAGCTGCTCATGAGTATTTATCA 85
OY 1961 CGCGCTGCTTTTATGATTCACAAATCTATTTTATTTGAGTTAAATAGAAAAA 2020
    |||||||
DB 84 CGCGCTGCTTTTATGATTCACAAATCTATTTTATTTGAGTTAAATAGAAAAA 25
OY 2021 ATTTACAGAAAAAATTTTAA 2041
    |||||||
DB 24 ATTTACAGAAAAAATTTTAA 4

RESULT 3
LOCUS BM847727 578 bp mRNA linear EST 06-MAR-2002
DEFINITION K-EST0127301 S13KMS5 Homo sapiens cDNA clone S13KMS5-50-D09 5',
            mRNA sequence.
ACCESSION BM847727.1 GI:19204126
VERSION EST.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 578)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            K.R.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
            21C Frontier Korean EST Project 2001
            Unpublished (2002)
COMMENT
JOURNAL
CONTACT: Kim YS
GENOME RESEARCH CENTER
KOREA RESEARCH INSTITUTE OF BIOSCIENCE & BIOTECHNOLOGY
52 BOEUN-DONG YUSEONG-GU, DAEJEON 305-333, SOUTH KOREA
TEL: +82-42-860-4409
FAX: +82-42-860-4470
EMAIL: yongsung@mail.kribb.re.kr
PLATE: 50 ROW: D COLUMN: 09
HIGH QUALITY SEQUENCE STOP: 578.
FEATURES
    source
        1. 578
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="S13KMS5-50-D09"
            /clone_lib="S13KMS5"
            /tissue_type="myeloma"
            /cell_line="KMS-5"
            /lab_host="Top10F"

```

```

BASE COUNT 150 a 143 c 160 g 125 t
ORIGIN
Query Match 26.5%; Score 541; DB 14; Length 578;
Best Local Similarity 100.0%; Pred. No. 2,5e-190;
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 44 GTGGCGAGTGGGCTCAGAGGCGCTGCAAGGACCCCTACCTCTGTAGTATTCGCGGGC 103
    |||||||
DB 38 GTGGCGAGTGGGCTCAGAGGCGCTGCAAGGACCCCTACCTCTGTAGTATTCGCGGGC 97
OY 104 ACGGGCACCGGCAATTCACGCGTGGGCTGACGTAAGCGACGCTCGCGGCTGAGATC 163
    |||||||
DB 98 ACGGGCACCGGCAATTCACGCGTGGGCTGACGTAAGCGACGCTCGCGGCTGAGATC 157
OY 164 GTGAGCGTACTCCATGACGAGGTCTATGAAGGCTTAGACATCATCCACAAGGTTCT 223
    |||||||
DB 158 GTGAGCGTACTCCATGACGAGGTCTATGAAGGCTTAGACATCATCCACAAGGTTCT 217
OY 224 GCCCAAGAGCAGAGAAATCTGCGGCGACACATGATAGCTTGTGATCTCTTGAGAC 283
    |||||||
DB 218 GCCCAAGAGCAGAGAAATCTGCGGCGACACATGATAGCTTGTGATCTCTTGAGAC 277
OY 284 AATTACACAGTGGTGAATGAGAAATGAGCACTGCTGTGATTAAGATATATTGGC 343
    |||||||
DB 278 AATTACACAGTGGTGAATGAGAAATGAGCACTGCTGTGATTAAGATATATTGGC 337
OY 344 CGAGACAAATTCCTATTGTTGTGGAGGAACCAATTATTACATTGAATCTGCTCTGG 403
    |||||||
DB 338 CGAGACAAATTCCTATTGTTGTGGAGGAACCAATTATTACATTGAATCTGCTCTGG 397
OY 404 AAAGTTCTTGTCAATACCAAGCCCGCAGAGATGGGACATGAGAAAGTATGACGAAAA 463
    |||||||
DB 398 AAAGTTCTTGTCAATACCAAGCCCGCAGAGATGGGACATGAGAAAGTATGACGAAAA 457
OY 464 GTGAGCTTGAAGAAGAGAGATGGTCTGTCTTACAAACGCTTAAGCCAGGTGACCCA 523
    |||||||
DB 458 GTGAGCTTGAAGAAGAGAGATGGTCTGTCTTACAAACGCTTAAGCCAGGTGACCCA 517
OY 524 GAAATGCTGCCAAGCTGCATCCACATGACAAACGCAAGTGGCCAGAGACTTGCAGATT 583
    |||||||
DB 518 GAAATGCTGCCAAGCTGCATCCACATGACAAACGCAAGTGGCCAGAGACTTGCAGATT 577
OY 584 T 584
    ||
DB 578 T 578

RESULT 4
LOCUS BE315223 725 bp mRNA linear EST 26-OCT-2000
DEFINITION 601141778F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3141643 5',
            mRNA sequence.
ACCESSION BE315223
VERSION BE315223.1 GI:9145717
KEYWORDS EST.
SOURCE human.

```

/note="Vector: pCNS; Site.1: EcoRI; Site.2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then deapped with tobacco acid pyrophosphatase (TAP). The deapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 725)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DMP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
Plate: LCM11. row: k. column: 20
High quality sequence stop: 720.

FEATURES
Source
1..725
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:3141643"
/clone_lib="NIH-MGC-9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCCAGG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 186 a 175 c 197 g 167 t

ORIGIN

Query Match 25.9%; Score 529; DB 10; Length 725;
Best Local Similarity 99.7%; Pred. No. 5.6e-186;
Matches 699; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

28 TGACGAGCAGTTCCTGCGGAGTGGGCTCAGGGGCTGCACAGCAGCCTTCTTGT 87
1 TGACAGCAGCAGTTCCTGCGGAGTGGGCTCAGGGGCTGCACAGCAGCCTTCTTGT 60
88 AGTATTTCTCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 147
61 AGTATTTCTCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
148 GCTCGGCGGCTGAGATGCTGAGCGCTGATCCATGCAAGGCTTATGAAGGCTTAGACATCAT 207
121 GCTCGGCGGCTGAGATGCTGAGCGCTGATCCATGCAAGGCTTATGAAGGCTTAGACATCAT 180
208 CACCAACAAGGTTTCTGCCAGAGAGAGATCTCCCGACACCATGATGACGTTTGT 267
181 CACCAACAAGGTTTCTGCCAGAGAGAGATCTCCCGACACCATGATGACGTTTGT 229
268 GATCCTCTTGAGCAATTTACAGAGTGGGCTTCAAAATAGAGCAACTGCTGTAT 327
240 GATCCTCTTGAGCAATTTACAGAGTGGGCTTCAAAATAGAGCAACTGCTGTAT 299
328 TGAAGATATATTTGCGCGAGCAAAATCTTATTTGTTGGGAGGAACCAATTTATACAT 387
300 TGAAGATATATTTGCGCGAGCAAAATCTTATTTGTTGGGAGGAACCAATTTATACAT 359
388 TGAATTTCTGCTCTGGAAGGTTCTTGTCAATACCAAGCCCGAGAGATGGGCACTGAGA 447
360 TGAATTTCTGCTCTGGAAGGTTCTTGTCAATACCAAGCCCGAGAGATGGGCACTGAGA 419
448 AGTATTTGAGCAAAAGTGGAGCTGGAAGAGAGATGCTTGTACTTCAAAAGCGCT 507
420 AGTATTTGAGCAAAAGTGGAGCTGGAAGAGAGATGCTTGTACTTCAAAAGCGCT 479
508 AAGCAGGTGAGCAAAAGTGGAGCTGGAAGAGAGATGCTTGTACTTCAAAAGCGCT 567

Db 480 AAGCAGGTGAGCAAAAGTGGAGCTGGAAGAGAGATGCTTGTACTTCAAAAGCGCT 539
Qy 568 CAGAGAGCTTGCAGATTTTGAAGAAGAGAGATGCTTGTACTTCAAAAGCGCT 627
Db 540 CAGAGAGCTTGCAGATTTTGAAGAAGAGAGATGCTTGTACTTCAAAAGCGCT 599
Qy 628 ACATACGAGAGAGAGATGCTTGCAGAGCTTGTCAAGTTCTTCAACCTTCAT 687
Db 600 ACATACGAGAGAGAGATGCTTGCAGAGCTTGTCAAGTTCTTCAACCTTCAT 659
Qy 688 CATTGCTTCATGCTGAGCAGGAGCTTGTATGAGCGCT 728
Db 660 CATTGCTTCATGCTGAGCAGGAGCTTGTATGAGCGCT 700

RESULT 5
AI207688/c 601 bp mRNA linear EST 11-NOV-1999
LOCUS HA3066 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
DEFINITION sequence.
ACCESSION AI207688
VERSION AI207688.1 GI:6361702
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 601)
AUTHORS Yu, Y., Zhang, C., Luo, L., Ouyang, S., Li, W., Wu, J., Zhou, S., Liu, M.
and He, F.
Expression profile analysis of a human fetal liver cDNA library
Unpublished (1998)
Contact: Chenggang Zhang
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Email: zhang_chenggang@hotmail.com.

FEATURES
Source
1..601
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human fetal liver cDNA library"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="MC1061/P3"
/note="Vector: pCDNA1"

BASE COUNT 198 a 137 c 102 g 164 t

ORIGIN

Query Match 23.5%; Score 480; DB 9; Length 601;
Best Local Similarity 99.8%; Pred. No. 9.5e-168;
Matches 600; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1389 AAGGATATGTTTGTCTCCAGCTGTCGCGCAAGAGAGCTATCGCGAATCTCTGCTAC 1448
601 AAGGATATGTTTGTCTCCAGCTGTCGCGCAAGAGAGCTATCGCGAATCTCTGCTAC 542
Qy 1449 AGAAAGCTCCACCATTTTCTTTGATGCTGTTTAAAGTCTCAAGCTCTCTATAAAG 1508
Db 541 AGAAAGCTCCACCATTTTCTTTGATGCTGTTTAAAGTCTCAAGCTCTCTATAAAG 482
Qy 1509 AAACAGCAGCTTGTGACCTCTTGTGCTGATGCTGGAAGATGATGATTCAG 1568
Db 481 AAACAGCAGCTTGTGACCTCTTGTGCTGATGCTGGAAGATGATGATTCAG 422
Qy 1569 GAAGCATTTTCTTTTGAACCTTAAGCTCTATTTAAAGCAGACAGATTG 1628
Db 421 GAAGCATTTTCTTTTGAACCTTAAGCTCTATTTAAAGCAGACAGATTG 362
Qy 1629 CACATTTTATACAGAGATCTTGTGCTGGAATACAGAGATTGATGATCCCTTT 1688
Db 361 CACATTTTATACAGAGATCTTGTGCTGGAATACAGAGATTGATGATCCCTTT 302

QY 1689 AAAAGAGTTTATGTCCTGACTGCTGCTAAATATCTAATTTCCAGATGCTTTGTA 1748
 |||||||
 Db 301 -AAGAAGATTATGTCCTGACTGCTGCTAAATATCTAATTTCCAGATGCTTTGTA 243
 |||||||
 QY 1749 GATGACTGAAGTATTTTGAGCCACATATTGGAGTTTGAATTTGATGAATGGCAGGA 1808
 |||||||
 Db 242 GATGACTGAAGTATTTTGAGCCACATATTGGAGTTTGAATTTGATGAATGGCAGGA 183
 |||||||
 QY 1809 AAGGCCATCTCCATTGATGATTAAGTGAACCAACTGTTCTGGAAATTTCTACAGAG 1868
 |||||||
 Db 182 AAGGCCATCTCCATTGATGATTAAGTGAACCAACTGTTCTGGAAATTTCTACAGAG 123
 |||||||
 QY 1869 AAGGCCATCTCCATTGATGATTAAGTGAACCAACTGTTCTGGAAATTTCTACAGAG 1928
 |||||||
 Db 122 AAGGCCATCTCCATTGATGATTAAGTGAACCAACTGTTCTGGAAATTTCTACAGAG 63
 |||||||
 QY 1929 TTTGCGAGCTGCTCATGTGAGTATTAATCACTGCTCTTTCTATTGATTAATC 1988
 |||||||
 Db 62 TTTGCGAGCTGCTCATGTGAGTATTAATCACTGCTCTTTCTATTGATTAATC 3
 |||||||
 QY 1989 T 1989
 |||||||
 Db 2 T 2

RESULT 6
 B1758117 799 bp mRNA linear EST 25-SEP-2001
 LOCUS 603023811F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5194347 5'
 DEFINITION mRNA sequence.
 ACCESSION B1758117
 VERSION B1758117.1 GI:15749695
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE NIH-MGC http://mgc.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: c9abps-rt@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.llnl.gov

plate: LAM1486 row: e column: 04
 High quality sequence start: 2
 High quality sequence stop: 775.
 Location/Qualifiers

FEATURES

1..799
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5194347"
 /clone_id="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH-MGC Library."

BASE COUNT 207 a 188 c 220 g 184 t
 ORIGIN

Query Match 22.8%; Score 466; DB 13; Length 799;

Best Local Similarity 99.6%; Pred. No. 1,1e-162;
 Matches 756; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 12 TGGGGTCGCGGGGGGCGACAGACAGTTCCTGAGGACATGAGGAGCCCTGCAAC 71
 |||||||
 Db 1 TGGCGTCGCGGGGGGCGACAGACAGTTCCTGAGGACATGAGGAGCCCTGCAAC 60
 |||||||
 QY 72 GGAACCTACCTCTTGTAGTATTCCTCGGGGCGACAGGAGCCGCGCAATTCACAGCTGGCGT 131
 |||||||
 Db 61 GGAACCTACCTCTTGTAGTATTCCTCGGGGCGACAGGAGCCGCGCAATTCACAGCTGGCGT 120
 |||||||
 QY 132 TGCAGTAGGCGACAGGCGCTCGGCGGTAGATCGTACAGCTGACTCCATGAGGTCTATG 191
 |||||||
 Db 121 TGCAGTAGGCGACAGGCGCTCGGCGGTAGATCGTACAGCTGACTCCATGAGGTCTATG 180
 |||||||
 QY 192 AAGGCGTACATCATATACCAACAAGTTTCTGCGCAAGAGCAGAGATCGCGCGAC 251
 |||||||
 Db 181 AAGGCGTACATCATATACCAACAAGTTTCTGCGCAAGAGCAGAGATCGCGCGAC 240
 |||||||
 QY 252 ACATGATCAGCTTTGTGATCCTCTTGTGACCAATTTACAGAGTGGAGCTTGAGAAATA 311
 |||||||
 Db 241 ACATGATCAGCTTTGTGATCCTCTTGTGACCAATTTACAGAGTGGAGCTTGAGAAATA 300
 |||||||
 QY 312 GAGCACTGCTCTGATTTGAGATATATTTGCGCGAGACAAATTTCTATTGTTGGGAG 371
 |||||||
 Db 301 GAGCACTGCTCTGATTTGAGATATATTTGCGCGAGACAAATTTCTA-TGTTGGGAG 359
 |||||||
 QY 372 GAACCAATTTATATGATGAAATCTGCTGCGGAAGTTCTTGCAATACCAAGCCCGAG 431
 |||||||
 Db 360 GAACCAATTTATATGATGAAATCTGCTGCGGAAGTTCTTGCAATACCAAGCCCGAG 419
 |||||||
 QY 432 AGATGGGACATGAGAAAGTGAATGACGAAAGTGAAGCTTGAAAGAGAGATGGTCTTG 491
 |||||||
 Db 420 AGATGGGACATGAGAAAGTGAATGACGAAAGTGAAGCTTGAAAGAGAGATGGTCTTG 479
 |||||||
 QY 492 TACTTCACAAAGCCTTAAGCCAGAGTGAACCCAGAAATGCTGCGCAAGCTGATCCATG 551
 |||||||
 Db 480 TACTTCAC -AAGCCTTAAGCCAGAGTGAACCCAGAAATGCTGCGCAAGCTGATCCATG 538
 |||||||
 QY 552 ACAAAAGCAAGTGGGCGGAGAGCTTGCAAGTTTGAAGAAACGAATCTCTATAGTG 611
 |||||||
 Db 539 ACAAAAGCAAGTGGGCGGAGAGCTTGCAAGTTTGAAGAAACGAATCTCTATAGTG 598
 |||||||
 QY 612 AATTCTCCATCGTCAACATACGGAAGAAGTGTGCTGCCCTTGAGAGTCTCTGAAGT 671
 |||||||
 Db 599 AATTCTCCATCGTCAACATACGGAAGAAGTGTGCTGCCCTTGAGAGTCTCTGAAGT 658
 |||||||
 QY 672 TCTCTAACCTTGCATCTCTTGGCTTATGCTGACACAGGAGTCTACATAGGCTTGG 731
 |||||||
 Db 659 TCTCTAACCTTGCATCTCTTGGCTTATGCTGACACAGGAGTCTAGTGAAGGCTTGG 718
 |||||||
 QY 732 ATAAGAGGAGTGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 770
 |||||||
 Db 719 ATAAGAGGAGTGAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 757
 |||||||

RESULT 7
 BMB00217 1088 bp mRNA linear EST 05-MAR-2002
 LOCUS AGNC00217 6416157 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5531248
 DEFINITION 5', mRNA sequence.
 ACCESSION BMB00217
 VERSION BMB00217.1 GI:19117040
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 1088)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

AUTHORS Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT

Qy	1	CTGGCCATTAGATAGGCGTCCGCGGGCGGCTGCACGACAGATCTCTGTAAGGACAGTGGGCTCAC	60
Db	5	CTGGCCATTAGATAGGCGTCCGCGGGCGGCTGCACGACAGATCTCTGTGGGCGATGGGCTCAG	64
Qy	61	GGGCGCTGCACAGGAGCCCTACCTCTTTGATGATTCTCGGGGCGCACGGGCGACCGGCAATC	120
Db	65	GGGCGCTGCACAGGAGCCCTACTCTTTGATGATTTCTCGGGGCGCACGGGCGACCGGCAATC	124
Qy	121	CACGCTGCAGTTGGCAGCTAGGCGCAGCGGCTGCGCGGTAGATCGTCAAGCGCTACTGTCAT	180
Db	125	CACGCTGCAGTTGGCAGCTAGGCGCAGCGGCTGCGCGGTAGATCTCTCAAGCGCTGACTGCAT	184

BASE COUNT	191 a	118 c	84 g	166 t
ORIGIN				
Query Match		19.0%;	Score 388;	DB 9, Length 559,
Best Local Similarity		99.6%;	Pred. No. 1.2e-133;	

	Matches	558:	Conservative	0:	Mismatches	1:	Indels	1:	Gaps	1:
QY	1470	TTTTATGTCGTTTAAAGTCACGTCCTATATATACAAACACGAGCTCTTGACGT								1529
Db	559	TTTTGATGTCGTTTAAAGTCACGTCCTCATATATAGAAACACGAGCTTGACGT								500
QY	1530	CCTTGTCGGCGATGTCGTGGAATGATGATGTCAGGAAGCATTTTTTTTCTTT								1589
Db	499	CCTTGTCGGCGATGTCGTGGAATGATGATGTCAGGAAGCATTTTTTTTCTTT								440
QY	1590	GAACCTTAAAGCTTCATTATTTAAAGCAGCAGACATTCACATTTTATACATGAGAT								1649
Db	439	GAACCTTAAAGCTTCATTATTTAAAGCAGCAGACATTCACATTTTATACATGAGAT								380
QY	1650	CTTCTTTTGTCGATACACGAGATTGACTGCATCCCTTTAAAGAAGTTTATGTCCTG								1709
Db	379	CTTCTTTTGTCGATACACGAGATTGACTGCATCCCTTT-AAAGAAGTTTATGTCCTG								321
QY	1710	ACTCGGCTAAATATCATANTTCCAGATGCTTTTGATAGACTGAAGTATTTCTGAG								1769
Db	320	ACTCGGCTAAATATCATANTTCCAGATGCTTTTGATAGACTGAAGTATTTCTGAG								261
QY	1770	CCACATATTGGGAGTTCTAGATTTCAGTGAATGGCAGAAAGGCCATCTCCATTGAGAT								1829
Db	260	CCACATATTGGGAGTTCTAGATTTCAGTGAATGGCAGAAAGGCCATCTCCATTGAGAT								201
QY	1830	GATTATAGTGAACCAACTAGTCTTCGCAATTCACAGAGAAAGGAGGAAATCAGACTGAGC								1889
Db	200	GATTATAGTGAACCAACTAGTCTTCGCAATTCACAGAGAAAGGAGGAAATCAGACTGAGC								141
QY	1890	AAGCGTCACATPAGACTTGAAGACCAAGACTTTGAAATTTGGCAGCTGCTCATCTGTG								1949
Db	140	AAGCTGTACACTAGCACTTGAAGACCAAGACTTTGAAATTTGGCAGCTGCTCATCTGTG								81
QY	1950	AGTTATATACACTGCTGTCCTTCTATTTAGGTATACAAATCTATATTTTATGACGTTTAA								2009
Db	80	AGTTATATACACTGCTGCTTCTATTTAGGTATACAAATCTATATTTTATGACGTTTAA								21
QY	2010	ATAAAGAAAAATTACAG 2029								
Db	20	ATAAAGAAAAATTACAG 1								

RESULT 11					
BQ003256/c					
LOCUS	BQ003256	624 bp	mRNA	linear	EST 26-MAR-2002
DEFINITION	U1-H-E11-ayx-n-12-O-U1.s1 NCI CGAP_E11 Homo sapiens				
	IMAGE:5845067 3', mRNA sequence.				

ACCESSION	B0003256	
VERSION	B0003256.1	GI:19728156
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 624)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.

Email: ccapds-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.S.E. Consortium/INL at: <http://image.inl.gov>
The following repetitive elements were found in this cDNA
sequence: 1-36, >POLY_A#Simple_repeat (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.

```

Source
1. 624
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:5845067"
/clone_lib="NCI CGAP E11"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Left Pelvis; Vector: pRTT3-Pac (Pharmacia) with a modified polylinker; Site:1: EcoR I; Site:2: Not I; NCI CGAP E11 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pRTT3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is
ACACTGCAC.
TAG_Lib-UI-H-E11
TAG_TISSUE-Chondrosarcoma
TAG_SEQ-ACACTGCAC"
203 a 128 c 97 g 195 t 1 others
BASE COUNT
ORIGIN

```

Query Match	19.0%	Score 387	DB 14	Length 624
Best Local Similarity	99.5%	Pred. 2.6e-133		
Matches 607	Conservative 0	Mismatches 23	Indels 1	Gaps 1
QY 1419	AGGAGTGCCTATGCGGAATTCCTCGATAGCAGCAAAAAGCTCCACCATTTCTTTTGATGT	1478		
DB 624	AGGAGTGCCTATGCGGAATTCCTCGATAGCAGCAAAAAGCTCCACCATTTCTTTTGATGT	565		
QY 1479	GSTTTAAAGTCTCACGTTCTCTATAATGAAGACAGAGTCTTGTCACTCCTGTGTG	1538		
DB 564	GSTTTAAAGTCTCACGTTCTCTATAATGAAGACAGAGTCTTGTCACTCCTGTGTG	505		
QY 1539	GCTGATGTGTCGGAATAGATGATAGTCAGAAACATTTTTTTCTTTGAACCTTAA	1598		
DB 504	GCTGATGTGTCGGAATAGATGATAGTCAGAAACATTTTTTTCTTTGAACCTTAA	445		
QY 1599	AGGTTCTATTATTAAAGCAGCAGATTCACATTTTAAATACATGAGATCTCTTGT	1658		
DB 444	AGGTTCTATTATTAAAGCAGCAGATTCACATTTTAAATACATGAGATCTCTTGT	385		
QY 1659	GCTGATATCCAGAGATGATGCTGCATCCCTTTAAAGATTAAATGTCCTGCTGTGCT	1718		
DB 384	GCTGATATCCAGAGATGATGCTGCATCCCTTTAAAGATTAAATGTCCTGCTGTGCT	326		
QY 1719	AAAAATTATCTAATTTCCAGATGCTTTTGTAGATGACTAGATTTGTGAGCACATATT	1778		
DB 325	AAAAATTATCTAATTTCCAGATGCTTTTGTAGATGACTAGATTTGTGAGCACATATT	266		
QY 1779	GGGAGTTCTAAGATTGATGAATGGCAGAGAAAGGCCATCTCCATTGAGATGATTAAGTG	1838		
DB 265	GGGAGTTCTAAGATTGATGAATGGCAGAGAAAGGCCATCTCCATTGAGATGATTAAGTG	206		
QY 1839	AACCAAACTAGTTCGCGAATTCCTACAGAGAAGAGGAATCAGACTAGCAAGCTGTGA	1898		
DB 205	AACCAAACTAGTTCGCGAATTCCTACAGAGAAGAGGAATCAGACTAGCAAGCTGTGA	146		
QY 1899	CATAGGACTTGAAGACCAAGACTTTGAAATTTGCGAGCTGCTATGTGTGAGTTATTAT	1958		
DB 145	CATAGGACTTGAAGACCAAGACTTTGAAATTTGCGAGCTGCTATGTGTGAGTTATTAT	86		
QY 1959	CACGCTCTCTCTTCTATGAGATTACAAATCATATTTTATTTGAAGTTTAAATAGAAA	2018		
DB 85	CACGCTCTCTCTTCTATGAGATTACAAATCATATTTTATTTGAAGTTTAAATAGAAA	26		
QY 2019	AAATTTACAA 2028			

Db	25	AAATTACAA	16
RESULT	12		
LOCUS	BI222788		
DEFINITION	60294157F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5104483 5'		
ACCESSION	BI222788		
VERSION	BI222788.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabds@email.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Incyte Genomics, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LHAM11252 row: d column: 20 High quality sequence stop: 741.		
FEATURES			
Source			
	1..783		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5104483"		
	/clone_lib="NIH_MGC_12"		
	/tissue_type="cervical carcinoma cell line"		
	/lab_host="DH10B"		
	/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 Kb. Library prepared by Life Technologies."		
BASE COUNT	218 a	181 c	184 g
ORIGIN			
	19.0%; Score 387; DB 13; Length 783;		
Query Match	Best Local Similarity 99.8%; Pred. No. 2.1e-133;		
Matches	507; Conservative	0; Mismatches	0; Indels
			1; Gaps
			1
QY	203	ATCATCACCACAAAGGTTTCTGCCCAAGACAGAGAATCTGCCGCGACACATGATCAGC	262
	1	ATCATCACCACAAAGGTTTCTGCCCAAGACAGAGAATCTGCCGCGACACATGATCAGC	60
Db	263	TTTGTGGATCTCTTGTGACCAATTACACAGTGTGGACTTTCGAATAGGCACTGCT	322
QY	61	TTTGTGGATCTCTTGTGACCAATTACACAGTGTGGACTTTCGAATAGGCACTGCT	120
Db	323	CTGATTTGAAGATATATTTTCCCGAGACAAATTCCTATTGTTGGAGGAACCAATTAT	382
QY	121	CTGATTTGAAGATATATTTTCCCGAGACAAATTCCTATTGTTGGAGGAACCAATTAT	180
Db	383	TACATTAATCTCTGCTCTGGAAGTTCTTGTTCATACCAAGCCCGAGAGATGGGACT	442
QY	181	TACATTAATCTCTGCTCTGGAAGTTCTTGTTCATACCAAGCCCGAGAGATGGGACT	240
Db	443	GAGAAAGTATTGACCGAAAGTGAGCTTGAA-AAAGAGATGGTCTTACTTACCA	501
QY	241	GAGAAAGTATTGACCGAAAGTGAGCTTGAA-AAAGAGATGGTCTTACTTACCA	300
Db	502	AGCCCTAAGCCAGGTGGAGCCAGAAATGGCTGCGCAAGCTGCATGCACATGACAAAGCAA	561
QY	301	AGCCCTAAGCCAGGTGGAGCCAGAAATGGCTGCGCAAGCTGCATGCACATGACAAAGCAA	560

OY	562	AGTGGCCAGAGACTTGCAACCTTTTGAAGAACAACGAAATGTCTCATAGTAATTTCICA	621
Db	361	AGTGGCCAGAGACTTGCAACCTTTTGAAGAACAACGAAATGTCTCATAGTAATTTCICA	420
OY	622	TTCGCACATACGGAAGAAGGTGGTGTCTCCCTTGAGAGCTCCTTGAAGTTCCTAACCC	681
Db	421	TTCGCACATACGGAAGAAGGTGGTGTCTCCCTTGAGAGCTCCTTGAAGTTCCTAACCC	480
OY	682	TTGCATCCTTTGGCTTCATGCTGACCAG	709
Db	481	TTGCATCCTTTGGCTTCATGCTGACCAG	508
RESULT 13			
LOCUS	AA121465		
DEFINITION	AA121465	457 bp	mRNA
ACCESSION	AA121465		linear
VERSION	AA121465.1		EST 19-NOV-1996
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	1 (bases 1 to 457)		
	Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman		
	,M., Hultman,M., Kucab,J., Le,M., Lennon,G., Marra,M., Parsons,J.,		
	Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston		
	,R., Williamson,A., Woldmann,P. and Wilson,R.		
	The Washu-Merck EST Project		
	Unpublished (1995)		
JOURNAL	Contact: Wilson RK		
COMMENT	Washington University School of Medicine		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
	Tel.: 314 286 1800		
	Fax: 314 286 1810		
	Email: est@watson.wustl.edu		
	This clone is available royalty-free through LNL ; contact the		
	IMAGE Consortium (info@image.llnl.gov) for further information.		
	Seq primer: -28M13 rev2 from AmerSham		
	High quality sequence stop: 419.		
FEATURES			
source	Location/Qualifiers		
	..457		
	/organism="Homo sapiens"		
	/db_xref="GDB:3804775"		
	/db_xref="taxon:9606"		
	/clone_image="IMAGE:490188"		
	/clone_lib="Soares_pregnant_uterus_NBHPU"		
	/sex="female"		
	/dev_stage="adult"		
	/lab_host="DH10B"		
	/note="Organ: uterus; Vector: pTZ19-Pac; Site_1: Not I ;		
	Site_2: Eco RI; 1st strand cDNA was primed with a Not I -		
	oligo(dT) primer [5'		
	AACGAGAGAAATCGCGCGCGCTTTTTTTTTTTTTTTTTTTT 3'] ,		
	double-stranded cDNA was ligated to Eco RI adaptors		
	(Pharmacia), digested with Not I and cloned into the Not I		
	and Eco RI sites of the modified pTZ19 vector. Library		
	went through one round of normalization. Library		
	constructed by M. Fatima Bernaldo."		
BASE COUNT	119 a	82 c	112 g
ORIGIN			142 t
			2 others
Query Match	18.9%	Score 385;	DB 9;
Best Local Similarity	99.8%;	Pred. No. 1.9e-133;	Length 457;
Matches 435;	Conservative	0;	Mismatches 1;
		Indels	0;
		Gaps	0;
OY	1287	GGAAGGATCCCCAGGCCGAGATGATCAAGAGCTGAATAATCGACCGTTTAAAGACATGTC	1346
Db	1	GGAAGGATCCCCAGGCCGAGATGATCAAGAGCTGAATAATCGACCGTTTAAAGACATGTC	60

QY 1347 CAGTGGCCTTTGGAAAGTGTGGGATCCAGTTACAGAGGAGGGTATGTTGCTCC 1405
 |||||||
 Db 61 CAGTGGCCTTTGCAAGAGTGTGGGATCCAGTTACAGAGGAGGGTATGTTGCTCC 120
 QY 1407 CAGTGGGCAAGAGAGTGTATGCGCAATTCCTGCATAGCAGAAAGCTCCACCATT 1466
 |||||||
 Db 121 CAGTGGGCAAGAGAGTGTATGCGCAATTCCTGCATAGCAGAAAGCTCCACCATT 180
 QY 1467 TTTCTTTGATGTGTTTAAAGTCTACGTTCTCTATATAGAAACAGAGTCTTGTC 1526
 |||||||
 Db 181 TTTCTTTGATGTGTTTAAAGTCTACGTTCTCTATATAGAAACAGAGTCTTGTC 240
 QY 1527 GCTCCTTGCTGCTGATGTGTGGAATGATGTTGAGAAAGCATTTTCTTTC 1586
 |||||||
 Db 241 GCTCCTTGCTGCTGATGTGTGGAATGATGTTGAGAAAGCATTTTCTTTC 300
 QY 1587 TTTGAACCTTAAAGTCTATATTAATAAAGCAGACAGATTCACATTTTATACATGAG 1646
 |||||||
 Db 301 TTTGAACCTTAAAGTCTATATTAATAAAGCAGACAGATTCACATTTTATACATGAG 360
 QY 1647 GATCTCTTTGCTGTAATACAGAGATGATGCTGATCCCTTAAAGAGTATTATGTC 1706
 |||||||
 Db 361 GATCTCTTTGCTGTAATACAGAGATGATGCTGATCCCTTAAAGAGTATTATGTC 420
 QY 1707 CTGACTCTGGCTAAA 1722
 |||||||
 Db 421 CTGACTCTGGCTAAA 436
 RESULT 14
 A1193562 543 bp mRNA linear EST 13-OCT-1998
 LOCUS A1193562/c
 DEFINITION g070f07.x1 Soares fetal_lung_NbHL19W Homo sapiens cDNA clone
 IMAGE:1744357 3', mRNA sequence.
 A1193562
 VERSION A1193562.1 GI:3744771
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 543)
 NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 Tumor Gene Index
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Unpublished (1997)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-remail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40UP from G1BCO
 High quality sequence stop: 472.
 FEATURES
 source
 1..543
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1744357"
 /clone_lib="Soares_fetal_lung_NbHL19W"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: Lung; Vector: p773D (Pharmacia) with a
 modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st
 strand cDNA was primed with a Not I - Oligo(dT) primer
 (5'-TGTTCACCAATCTGAGTGGAGCGCCGCAATTTTCTTTTCTTTT-3'),
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified p773 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M. Paloma Bonaldo. This library was constructed
 from the same fetus as the fetal heart library, Soares
 fetal heart NbHL19W."
 BASE COUNT 181 a 114 c 84 g 164 t

ORIGIN
 Query Match 18.2%; Score 372; DB 9; Length 543;
 Best Local Similarity 99.6%; Pred. No. 1,1e-127;
 Matches 542; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1487 AGTCTACGTTCTCTATATAGAAACAGAGCTTTGTCAGCTCTTTGCTGCTATGCT 1546
 |||||||
 Db 543 AGTCTACGTTCTCTATATAGAAACAGAGCTTTGTCAGCTCTTTGCTGCTATGCT 484
 QY 1547 GCTCGAAATGATGTTGTTGAGAAAGCATTTTCTTTCCTTGAACCTTAAAGTTCT 1606
 |||||||
 Db 483 GCTCGAAATGATGTTGTTGAGAAAGCATTTTCTTTCCTTGAACCTTAAAGTTCT 424
 QY 1607 TTATTAAGCAGACAGATTCACATTTTATACATGAGAGATCTTCTTGTGTAATA 1666
 |||||||
 Db 423 TTATTAAGCAGACAGATTCACATTTTATACATGAGAGATCTTCTTGTGTAATA 364
 QY 1667 CCAGAGTGAATGATGCTTTTAAAGAGTTTATGTCCTGCTGCTGCTAAATAT 1726
 |||||||
 Db 363 CCAGAGTGAATGATGCTTTTAAAGAGTTTATGTCCTGCTGCTGCTAAATAT 305
 QY 1727 CTATTTCCAGATGCTTTTGTAGATGATGATGATGATGATGATGATGATGATGAT 1786
 |||||||
 Db 304 CTATTTCCAGATGCTTTTGTAGATGATGATGATGATGATGATGATGATGATGAT 245
 QY 1787 TAGATTGAGTGAATGAGCAAGAAAGGCCATCTCATTTAGATGATTAAGTGAACCAAC 1846
 |||||||
 Db 244 TAGATTGAGTGAATGAGCAAGAAAGGCCATCTCATTTAGATGATTAAGTGAACCAAC 185
 QY 1847 TAGTTCTCGAATCTTACAGAGAGAGGAATACACAGCAGAGAGCTGTGATGAGAC 1906
 |||||||
 Db 184 TAGTTCTCGAATCTTACAGAGAGAGGAATACACAGCAGAGAGCTGTGATGAGAC 125
 QY 1907 TTGAAGACCAAGACTTTGAAATTTGCGAGCTGCTCATGCTGAGATTATATACATGAG 1966
 |||||||
 Db 124 TTGAAGACCAAGACTTTGAAATTTGCGAGCTGCTCATGCTGAGATTATATACATGAG 65
 QY 1967 TCTTTCTATTGAGTACCAATCTATATTTTATGAAGTTTAAATAAGAAAAAATTAC 2026
 |||||||
 Db 64 TCTTTCTATTGAGTACCAATCTATATTTTATGAAGTTTAAATAAGAAAAAATTAC 5
 QY 2027 AAGA 2030
 |||||
 Db 4 AAGA 1
 RESULT 15
 BG495857 562 bp mRNA linear EST 27-MAR-2001
 LOCUS BG495857
 DEFINITION 602540248P1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4671586 5',
 mRNA sequence.
 ACCESION BG495857
 VERSION BG495857.1 GI:13457373
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 562)
 NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
 Tissue Procurement: ATCC
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LCM1486 row: 0 column: 11

FEATURES High quality sequence stop: 562.
Location/Qualifiers
source 1. 562
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4671586"
/clone_id="NIH_MGC_59"
/tissue_type="mucoepidermoid carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site: 1:
Site1 (99ccgctcgcc); Site2: Site1 (99ccattatgccc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATCTAGAGCCGCGCCGCGCATG-dT(30)BN-3'
sequence: 5'-ATCTAGAGCCGCGCCGCGCATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT 167 a 122 c 142 g 131 t
ORIGIN
Query Match 18.2% Score 371; DB 12; Length 562.
Best Local Similarity 99.4% Pred. No. 2.5e-127;
Matches 521: Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 939 GACCTGTCCATTCTCCCTCTGTCTATGCTTAGAGTATCTGATCTCTGAAGTGGG 998
DB 23 GACCTGTCCATTCTCCCTCTGTCTATGCTTAGAGTATCTGATCTCTGAAGTGGG 82
OY 999 AGGACTCTGTCTTGAACCTGCTTGAATCTGCAAGTTTCTCCAGGGCCACAAGC 1058
DB 83 AAGACTCTGTCTTGAACCTGCTTGAATCTGCAAGTTTCTCCAGGGCCACAAGC 142
OY 1059 CTACAGCCACTCCATTAAGATGGCATCAATGAAGCTGAGACAGAAGAACTTATCACC 1118
DB 143 CTACAGCCACTCCATTAAGATGGCATCAATGAAGCTGAGACAGAAGAACTTATCACC 202
OY 1119 TGTGTGACCTCTGTGATGAATCATCTGGGATCGGCAATGGCGCACATAAAT 1178
DB 203 TGTGTGACCTCTGTGATGAATCATCTGGGATCGGCAATGGCGCACATAAAT 262
OY 1179 CCAATCCCACTTGAACCACTGAAGAAGAAAGATGGACTCAGATGCTGCACAA 1238
DB 263 CCAATCCCACTTGAACCACTGAAGAAGAAAGATGGACTCAGATGCTGCACAA 322
OY 1239 CCATGAAGAAGTGAAGTGTTCCTCCAGACTATACAAAGAACCTAAAGGAGGATCCC 1298
DB 323 CCATGAAGAAGTGAAGTGTTCCTCCAGACTATACAAAGAACCTAAAGGAGGATCCC 382
OY 1299 CAGGCGAGATGATCAAGAGCTGAAGTCAAGCGTTTAAGAGACATGTCAGTGGCTTTG 1358
DB 383 CAGGCGAGATGATCAAGAGCTGAAGTCAAGCGTTTAAGAGACATGTCAGTGGCTTTG 442
OY 1359 GAAAGTGTGGGATCCAGTTCAGAGGAGAGGGTATGTTGCTCCAGTCTGGGCAG 1418
DB 443 GAAAGTGTGGGATCCAGTTCAGAGGAGAGGGTATGTTGCTCCAGTCTGGGCAG 502
OY 1419 AGGAGTGTATGCGGAATTTCTGATAGCAGAAAGCTCCAC 1462
DB 503 AGGAGTGTATGCGGAATTTCTGATAGCAGAAAGCTCCAC 546

Search completed: April 21, 2003, 22:17:05
Job time : 3302 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 18:47:50 ; Search time 95.2977 Seconds
(without alignments)
8825.857 Million cell updates/sec

Title: US-09-513-151-3
Perfect score: 3575
Sequence: 1 CTCGCCATAGATCGCTCCG.....TTTACAGAGAAAAAAA 2041

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+np.model -DEV=x1h
-Q=/cgn2_1/USPTO.spool/US09513151/runatc_15042003_141144_26380/app-query.fasta_1.2446
-DB=SPREMBL_21 -QFMT=fastan -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun2 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09513151.qcgn_1.1_125-efunat_15042003_141144_26380 -NCPu=6 -ICFu=3
-NO_XLPY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WAIN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: SPREMBL_21.*
2: sp_archaea.*
3: sp_bacteria.*
4: sp_fungi.*
5: sp_human.*
6: sp_invertebrate.*
7: sp_mammal.*
8: sp_mhc.*
9: sp_organelle.*
10: sp_phase.*
11: sp_plant.*
12: sp_rodent.*
13: sp_virus.*
14: sp_vertebrate.*
15: sp_unclassified.*
16: sp_rvivirus.*
17: sp_bacteriap.*
18: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query Score Match Length DB ID Description
1 2256 63.1 435 4 096L45 096L45 homo sapien

2	2253	63.0	467	4	09H3H1	09h3h1 homo sapien
3	1550	43.4	326	4	09NXT7	09nxt7 homo sapien
4	1531	42.8	324	4	096FJ3	096fj3 homo sapien
5	1342	37.5	326	11	09D1H5	09d1h5 mus musculu
6	549	15.4	430	5	09G7G3	09g7g3 caenorhabdi
7	528.5	14.8	434	3	09U775	09u775 schistosach
8	493	13.8	466	10	09ZUX7	09zux7 arabidopsis
9	482.5	13.5	439	10	09S9Z6	09s9z6 oryza sativ
10	438.5	12.3	315	16	09R5S5	09r5s5 thermoaer
11	339	9.5	427	10	09S9W3	09s9w3 oryza sativ
12	338	9.5	318	10	09S6B0	09s6b0 arabidopsis
13	334.5	9.4	329	10	0941D1	0941d1 arabidopsis
14	333.5	9.3	329	10	09LUG4	09lug4 arabidopsis
15	328.5	9.2	330	10	0941D2	0941d2 arabidopsis
16	327.5	9.2	330	10	0941C9	0941c9 arabidopsis
17	326	9.1	330	10	0941L4	0941l4 arabidopsis
18	322.5	9.0	357	10	09CA35	09ca35 arabidopsis
19	321.5	9.0	357	10	0941D3	0941d3 arabidopsis
20	313.5	8.8	330	10	09LVB1	09lvb1 arabidopsis
21	313.5	8.8	336	10	093WC9	093wc9 arabidopsis
22	302	8.4	463	10	09C5J6	09c5j6 arabidopsis
23	300	8.4	459	10	0941C8	0941c8 arabidopsis
24	299	8.4	350	10	09S533	09s533 petunia hyb
25	283	7.9	303	16	09R5Z6	09r5z6 fusobacteri
26	276.5	7.7	379	5	08TJL1	08tjl1 dictyostell
27	267	7.5	342	10	09C6L1	09c6l1 arabidopsis
28	239.5	6.7	297	2	08VOS5	08vos5 methylobact
29	234	6.5	330	10	09RXY9	09rxy9 arabidopsis
30	135.5	3.8	706	10	09CGR8	09cgr8 arabidopsis
31	129	3.6	982	16	08UF86	08uf86 agrobacteri
32	128.5	3.6	820	16	09WY41	09wy41 thermotoga
33	125	3.5	447	17	08TUY9	08tuy9 methanopyru
34	117.5	3.3	559	16	099RX4	099rx4 staphylococ
35	117	3.3	383	5	09VGR9	09vgr9 dtrosophila
36	117	3.3	363	16	092BY7	092by7 listeria in
37	116.5	3.3	952	10	09SXB7	09sxb7 arabidopsis
38	116.5	3.3	952	10	09M2Z6	09m2z6 arabidopsis
39	116	3.2	5171	4	08WKK9	08wkk9 homo sapien
40	116	3.2	26926	4	010466	010466 homo sapien
41	116	3.2	26926	4	08W2B3	08w2b3 homo sapien
42	116	3.2	34350	4	08W242	08w242 homo sapien
43	115	3.2	325	16	08X151	08x151 clostridium
44	113	3.2	763	10	09LM99	09lm99 arabidopsis
45	112	3.1	747	2	053832	053832 staphylococ

ALIGNMENTS

RESULT 1
096L45 PRELIMINARY: PRT: 435 AA.
AC 096L45.
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE tRNA isopentenyl transferase (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxId:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2444833; PubMed=11560893;
RA Lemaux J., Lakowski B., Webb A., Meng Y., Ubach A., Bussiere F.,
RT "Regulation of Physiological Rates in Caenorhabditis elegans by a
RT tRNA-Modifying Enzyme in the Mitochondria";
RL Genetics 159:147-157(2001).
DR EMBL: AY052268; AL14107.1; -
DR InterPro: IPR002627; IPTT:
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF01715; IPTT: 1.
DR ProDom: PD004674; IPTT: 1.

DR SMART; SM00355; ZnF_C2H2_1.
 DR TIGR00174; miaa; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 KW Transferase.
 FT NON-TER
 SQ SEQUENCE 435 AA: 48948 MW: 2279AE7C2D999FE1 CRC64;

Alignment Scores:

Pred. No.:	2,15e-180	Length:	435
Score:	2256.00	Matches:	424
Percent Similarity:	99.77%	Conservative:	0
Best Local Similarity:	99.77%	Mismatches:	1
Query Match:	63.10%	Indels:	0
DB:	4	Gaps:	0

US-09-513-151-3 (1-2041) x Q96L43 (1-435)

```

OY 29 GCACGAGAGTTCTGTGGGACATGGGCTCAGGGGCTGCACAGGACCTACCTTTGTA 88
DB 1 AlaArgAlaValProValGlySerGlyLeuArgGlyLeuGlnArgThrLeuProLeuVal 20
OY 89 GTGATTTCTGGGGCCAGGGGACCGGCAATCCAGCTGGCGTTCAGCTAGGCCAGCG 148
DB 21 ValIleLeuGlyAlaThrGlyThrGlyLysSerThrLeuAlaLeuGlnLeuGlyGlnArg 40
OY 149 CTCGGCGGTGAGATGCTCAAGCGCTGACTCCATGCAAGCTATGTAAGGCGCTAGCATCATC 208
DB 41 LeuGlyGlyLeuIleValSerAlaAspSerMetGlnValTyrGlnLeuAspIleIle 60
OY 209 ACCACAAAGTTTCTGCCAGAGAGCAGAGAAATCGCCGCGCACACATGATCAGCTTTGTC 268
DB 61 ThrAsnLysValSerAlaGlnGlnGlnArgIleCysArgHisHisMetLysSerPheVal 80
OY 269 GATCCCTCTGTGACCAATACACAGTGGTGGACTCAGAAATAGACCAACTGCTGTGATT 328
DB 81 AspProLeuValThrAsnTyrThrValValAspPheArgAsnArgAlaThrAlaLeuIle 100
OY 329 GAAGATATATTGGCCGAGACAAATCTCTATGTGTGGAGAGACCAATTATACATT 388
DB 101 GluAspIlePheAlaArgAspLysIleProIleValValGlyLysThrAsnTyrTyrIle 120
OY 389 GAATCTCTGCTGGAAGCTTTCTTGAATACCAAGCCCCAGAGATGGGCACTGAGAAA 448
DB 121 GluSerLeuLeuTyrLysValLeuValAsnThrLysProGlnGlnMetGlyThrGlyLys 140
OY 449 GTGATGACCGCAAAAGTGGAGCTTGAAGAAGAGATGGTGTGTACTTACACAACCCCTA 508
DB 141 ValIleAspArgLysValIleLeuGlnLysGlnAspGlyLeuValLeuHisLysArgLeu 160
OY 509 AGCAGAGTGGACCCAGAAATGGCTGCCAGCTGCATCCATGACAAAGCAAGTGGCC 568
DB 161 SerGlnValAspProGlnMetAlaAlaLysLeuHisProHisAspLysArgLysValAla 180
OY 569 AGCAGCTTGCAGATTGTTGAAGAAACAGAAATCTCATAGTAATTTCTCCATCTGCA 628
DB 181 ArgSerLeuGlnValPheGlnGlnThrGlyLysSerHisSerGlnPheLeuHisArgGln 200
OY 629 CATACGGAAGAGTGGTGGCTCCCTGAGAGTCTCTGAAGTCTCTTAACCTTGCATC 688
DB 201 HisThrGlnGlnGlyGlyGlyProLeuGlnGlyProLeuLysPheSerAsnProCysIle 220
OY 689 CTTTGGCTTCATGCTGACAGCAGCTTGTAGATGAGCGGTGATAGAGGGGTGAC 748
DB 221 LeuTyrPheHisAlaAspGlnAlaValAlaLeuAspGlnAlaValLeuAspArgLysArgValAspAsp 240
OY 749 ATGCTTGTGCTGGGCTCTTGGAGGAATAGAGATTTTACAGAGCGCTATATATCAGAG 808
DB 241 MetLeuAlaAlaGlyLeuLeuGlnGlnLeuLeuArgAspPheHisArgArgLysAsnGlnLys 260
OY 809 AATGTTTCCGAATAAGCCAGAGACTATCAACATGCTATATCTTCAATCATTTGGCTTCAAG 868
DB 261 AsnValSerGlnAsnSerGlnAspTyrGlnHisGlyIlePheGlnSerIleGlyPheLys 280

```

```

OY 869 GAATTTCACGAGTACCTGATTCACCTGAGGAAAAATGCACACTGAGACTAACCAGCTT 928
DB 281 GluPheHisGlyLysThrLeuIleThrGlnGlyLysCysThrLeuGlnThrSerAsnGlnLeu 300
OY 929 CTAAGAAGAGACCTGCTCCATTTGTCGCCCTGTCTATGCTTAAAGCTATCTGATGTC 988
DB 301 LeuLysLysGlyProGlyProIleValProProValTyrGlyLeuGlnValSerAspVal 320
OY 989 TCGAAGTGGAGAGAGTCTGTCTTGAACCTGCTTGAATTCGCAAGTTTCAATCCAG 1048
DB 321 SerLysTyrGlnGlnLysSerValLeuGlnProAlaLeuGlnIleValGlnSerPheIleGln 340
OY 1049 GGGCCACAGCCTACAGCCGCTCCATTAAGATGGCCATACAAAGAACCTGAGAACAGAGA 1108
DB 341 GlyHisLysProThrAlaThrProIleLysMetProTyrAsnGlnAlaGlnLysLysArg 360
OY 1109 AGTTATCACCTGTGTGACCTCTGTGATTCGAATCATCATTTGGGATTCGCAATGGCAGCG 1168
DB 361 SerTyrHisLeuGlnCysAspLeuGlnCysAspArgIleIleIleGlyAspArgGlnTyrPheAla 380
OY 1169 CACATTAATTCCAATCCCACTTGAACCAACTGAGAGAAAGAAAGAGATTGCACTCAGAT 1228
DB 381 HisIleLysSerLysSerHisLysAsnGlnLeuLysLysArgArgLysAspSerAsp 400
OY 1229 GCTGTCAACACATAGAAAGTCAAGTGTTCCTCCGACTATTAACAAGACCTTAAGGG 1288
DB 401 AlaValAsnThrIleGlnLysSerGlnSerValSerProAspTyrAsnLysGlnProLysGln 420
OY 1289 AAGGATCCCGCAGGCGAGATGATCAAGAGCTGAATTCAGCGGTT 1333
DB 421 LysGlySerProGlyLysAsnAspGlnGlnLeuLysCysSerVal 435

RESULT 2
ID Q9H3H1 PRELIMINARY; PRT; 467 AA.
AC Q9H3H1;
DT 01-MAR-2001 (TRMBLrel. 16, Created)
DT 01-MAR-2001 (TRMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TRMBLrel. 21, Last annotation update)
DE tRNA isopentenyl pyrophosphate transferase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20564178; PubMed=11111046;
RA Golovko A., Hjalms G., Silbon F., Nicander B.;
RT "Cloning of a human tRNA isopentenyl transferase."
RL Gene 258:85-93(2000).
DR EMBL: AF074918; AAC31324.1; -.
DR InterPro: IPR002627; IPT.
DR InterPro: IPR000822; ZnF_C2H2.
DR Pfam: PF01715; IPT. 1.
DR ProDom: PD004674; IPT. 1.
DR TIGR00174; miaa; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 467 AA: 52725 MW: 63446991D7F56A5 CRC64;

Alignment Scores:
Pred. No.: 3,91e-180 Length: 467
Score: 2253.00 Matches: 439
Percent Similarity: 94.22% Conservative: 1
Best Local Similarity: 94.00% Mismatches: 1
Query Match: 63.02% Indels: 26
DB: 4 Gaps: 1

US-09-513-151-3 (1-2041) x Q9H3H1 (1-467)
OY 11 ATGGCGTCCGTCGCGCTGCAGACAGATTCTGTGGCAGTGGGCTCAGGGGCTGCGAA 70
DB 1 MetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArgGlyLeuGln 20

```

```

OY 71 CGACCCCTACCTCTGTAGTATCTCGGGGCGACGCGCAACGCAATCCACGCTGGC 130
    |||||||
DB 21 ArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThcGlyLysSerThrLeuAla 40
OY 131 TTGCAGCTAGGCGCAGCGGCTCGGCGGTAGATGTCAGCGCTGATCCATGCAAGCTTAT 190
    |||||||
DB 41 LeuGlnLeuGlyGlnArgLeuGlyGlyGlnIleValSerAlaAspSerMetGlnValTyr 60
OY 191 CAAGCGCTAGACATCATACCAACAGGTTTGTGCCCAAGACAGACAGAAATCTCCGGCAG 250
    |||||||
DB 61 GlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIleCysArgHis 80
OY 251 CACATGATCAGCTTGTGATCCCTCTGTGACCAATTTACACAGTGTGATCCAGTTCAGAAAT 310
    |||||||
DB 81 HisMetIleSerPheValAspProLeuValThrAsnTyrThrValValAspPheArgAsn 100
OY 311 AGAGCACTGCTGTATGATGAAGATATATTTGCCCGAGACAAATTCCTATTGTTGGGA 370
    |||||||
DB 101 ArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIleValValGly 120
OY 371 GGAACCAATTTATACATTTGAATCTGCTGTGAAAGTTCTTGTCAATACCAAGCCCGAG 430
    |||||||
DB 121 GlyThrAsnTyrTyrIleGlnSerLeuLeuTyrPlysValLeuValAsnThrLysProGln 140
OY 431 GAGATGGCGCACTGAGAAAGTATGACCGAAAGTGGACCTGAAAGAGGATGGCTTT 490
    |||||||
DB 141 GlnMetGlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysLysArgLysLeu 160
OY 491 GTACTTTCACAAAGCGCTAAGCGAGTGGACCCAGAAATGGTTCGCAAGCTGCATCCACAT 550
    |||||||
DB 161 ValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaAlaLysLeuHisProHis 180
OY 551 GACAAACGCAAAAGTGGCGCAGAGCTTTCGCAAGTTTGAAGAAACAGAAATCTCTCATGT 610
    |||||||
DB 181 AspLysArgLysValAlaIleArgSerLeuGlnValPheGlnGlnThrGlyIleSerHisSer 200
OY 611 GAATTTCTCCATCGTCAACATACGAGAGAGAGTGGTGGCTTGGAGCTGCTCGTGAAG 670
    |||||||
DB 201 GlnPheLeuHisArgGlnIleThrGlnGlnGlyGlyGlyProLeuGlyLysProLeuLys 220
OY 671 TTCTSTAACCTTTGATCCTTTGGCTTCACTGCTGACAGGAGCTTCTAGATGAGCGCTTGG 730
    |||||||
DB 221 PheSerAsnProCysIleLeuTyrPheHisAlaAspGlnAlaValLeuAspGlnArgLeu 240
OY 731 GATTAAGAGGCTGATGACATGCTGCTGCTGCTGCTTGTGAGAGACTAAGATTTTTCAC 790
    |||||||
DB 241 AspLysArgValAspAspMetLeuAlaIleGlyLeuLeuGlnGlnIleuArgAspPheHis 260
OY 791 AGACCGTATTAATCAGAGAAATGTTTCGAAATATAGCCAGACTATCAACATGATGATCTTC 850
    |||||||
DB 261 ArgAlaGlyTyrAsnGlnLysAsnValSerGlnAsnSerGlnAspTyrGlnIleGlyIlePhe 280
OY 851 CAATCAATTTGCTTCAAGCAATTTTCAAGCACTGATCACTGAGGAGAAATTCACACTG 910
    |||||||
DB 281 GlnSerIleGlyPheLeuGlnPheHisGlnTyrLeuIleThrGlnGlnLysCysThrLeu 300
OY 911 GAGACTAGTAAACAGCTTTTAAAGAAAGG----- 940
    |||||||
DB 301 GlnThrSerAsnGlnLeuLeuLysGlyIleGlnAlaLeuLysGlnValThrLysArg 320
OY 941 -----CCTGGTCCCAT 952
    |||||||
DB 321 TyrAlaValGlnAsnArgTyrValLysAsnArgPheLeuSerArgProGlnProIle 340
OY 953 GTCCCGCTGTATGCTTGAAGATATGATGCTGCTGCAAGTGGAGGAGCTGTATTTT 1012
    |||||||
DB 341 ValProGlnValTyrGlyLeuGlnValSerAspValSerLysTyrPrlGlnSerValLeu 360
OY 1013 GAACGCTGCTTGAATGCTGCAAGTTCATGTCAGGCGCAGACGCTACAGCAGCTGCA 1072
    |||||||
DB 361 GlnProAlaLeuGlnIleValGlnSerPheIleGlnGlnLysLysProThrAlaThrPro 380

```

```

OY 1073 ATTAAGATGCCATACATGAAGCTGAGAACAGAGAACTTATACCTGTGACCTGTCT 1132
    |||||||
DB 381 IleLysMetProTyrAsnGlnIleAspGlnLysAsnLysArgSerTyrHisLeuCysAspLeuLys 400
OY 1133 GATCGAATCATCATTTGGGATGGCGAATGGCGACGCGACATATAATCCAAATCCCACTTGT 1192
    |||||||
DB 401 AspArgIleIleIleGlyAspArgGlnTyrPrlAlaAlaHisIleLysSerLysSerHisLeu 420
OY 1193 AACCAACTGAGAAAGAAAGAAAGATTGGATGCGATGCTGTCACACCATACAGAAAGTCAG 1252
    |||||||
DB 421 AsnGlnLeuLysLysArgArgArgLysAspSerAspAlaValAsnThrIleGlnSerGln 440
OY 1253 AGCTTTTCCCGCACTATTAACAAAGACCTAAAGGAGAGATCCCGACGCGAGATGAT 1312
    |||||||
DB 441 SerValSerProAspHisAsnLysGlnPrlGlnLysGlySerProGlnLysAsnAsp 460
OY 1313 CAAGACGCTGAATGCAGCTT 1333
    |||||||
DB 461 GlnGlnLeuLysCysSerVal 467

```

RESULT 3

```

O9NXT7 PRELIMINARY: PRT: 326 AA.
ID O9NXT7
AC O9NXT7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CDNA FLJ20061 f1s, clone COL01383.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obaishi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AK000068; BAA90923.1;
DR InterPro: IPR002627; IPTT.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF01715; IPTT. 1.
DR ProDom: PD004674; IPTT. 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2.1; UNKNOWN.1
SQ SEQUENCE 326 AA; 37435 MW; EAB3F0P9664B7ACE CRC64;

```

Alignment Scores:

```

Pred. No.: 2,84e-121 Length: 326
Score: 1550.00 Matches: 298
Percent Similarity: 91.728 Conservative: 1
Best Local Similarity: 91.418 Mismatches: 1
Query Match: 43.368 Indels: 26
DB: Gaps: 1

```

US-09-513-151-3 (1-2041) x O9NXT7 (1-326)

```

OY 434 ATGGGCACTGAGAAATGATTTGACCGAAAGTGGACTTGAAGAGAGATGCTGTGA 493
    |||||||
DB 1 MetGlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysAspIleLeuVal 20
OY 494 CTTCAACAAAGCTTGAAGCGAGGTGGACGAGCAAGAAATGGTGGCAAGCTGATCCATGAC 553
    |||||||
DB 21 LeuHisLysArgLeuSerGlnValAspProGlnMetAlaAlaLysLeuHisProHisAsp 40
OY 554 AAAGCAAGTGGCGCAGAGCTTTCAGATTTTGAAGAAGAGAAATGCTGTGATGTA 613
    |||||||
DB 41 LysArgLysValAlaIleArgSerLeuGlnValPheGlnGlnThrGlyIleSerHisSerGln 60
OY 614 TTTTCTCATGCTCAACATAGGAGAAAGTGGTGGCTGCTGCTGCAAGCTTGTGAAGTTC 673
    |||||||
DB 61 PheLeuHisArgGlnHisThrGlnGlnGlyGlyGlyProLeuGlnGlyProLeuLysPhe 80

```

OY 674 TTAACCCCTTCATCCCTTGGCTTCATGTCGACAGGACGCTTCAGATGACGCTTCGAT 733
 DB 81 SerAsnProCysIleLeuTrpLeuHisAlaAspIleAlaValLeuAspGluArgLeuAsp 100
 OY 734 AAGAGGCTGATGACATGCTTGGCTGCTGGGCTTCGAGAGAACTAAAGATTTTTCACAGA 793
 DB 101 LysArgValAspAspMetLeuAlaIleGlyLeuLeuGluIleuAlaGAspPheHisArg 120
 OY 794 CGCTATATATACAGAAAGATGTTTCGGAATAATAGCCAGACTATACATGCTATCTTCAA 853
 DB 121 ArgTyrAsnGlnLysAsnValSerGluAsnSerGlnAspTyrGlnHisGlyIlePheGln 140
 OY 854 TCATGCTGCTCAAGAAATTCACGAGTACCTGATCATCTGAGGAAATACACACTGAG 913
 DB 141 SerIleGlyPheLysGluPheHisGlyTyrLeuIleThrGluGlyLysCysThrLeuGlu 160
 OY 914 ACTAGTAAACGCTTCTTAAGAAAGCA-----CCTGCTCCATTGTC 940
 DB 161 ThrSerAsnGlnLeuLeuLysGlyIleGluAlaLeuLysGlnValThrLysArgTyr 180
 OY 941 -----CCTGCTCCATTGTC 955
 DB 181 AlaArgLysGlnAsnArgTyrValLysAsnArgPheLeuSerArgProGlyProIleVal 200
 OY 956 CCCCCTGTATAGGCTTAGAGATGATGATGCTCGAAGTGGAGGAGTCTGTTCTTGA 1015
 DB 201 ProProValIlyrGlyLeuGluValSerAspValSerLysTrpGluGlnSerValLeuGlu 220
 OY 1016 CCTGCTCTTGAATCGTGAAGAGTTTCATCCAGGCGCACAAAGCTTACAGCCACTCCATA 1075
 DB 221 ProAlaLeuGluIleValGlnSerPheIleGlnGlyHisLysProThrAlaThrProIle 240
 OY 1076 AAGATGCTATACATGATGAGTGAAGACAGACAGAAAGTATGATACCTGTGACTGTGAT 1135
 DB 241 LysMetProTyrAsnGlnAlaGlnAsnLysArgSerTyrHisLeuLysAspLeuLysAsp 260
 OY 1136 CGATCATCATGATGGGATCCGCAATGGGCGCACATAAATCCAAATCCCACTTGAAC 1195
 DB 261 ArgIleIleIleIleGlyAspArgGluTTPAlaIleHisIleLysSerLysSerHisLeuAsn 280
 OY 1196 CAACTGAAGAAAGAGAAAGATTGAGACTGATGCTGTCAACACATAGAAAGTCAAGT 1255
 DB 281 GlnLeuLysLysArgArgArgLeuAspSerAspAlaValAsnThrIleGlnSerGlnSer 300
 OY 1256 GTTCCCGACACTATTAACAAGAACTAAAGGAGGATCCCGAGGCGCAATGATCAA 1315
 DB 301 ValSerProAspHisAsnLysGluProLysGluLysGlySerProGlyGlnAsnAspGln 320
 OY 1316 GAGCTGAATGACGCGTT 1333
 DB 321 GlnLeuLysCysSerVal 326
 RESULT 4
 096FJ3 PRELIMINARY: PRT: 324 AA.
 AC 096FJ3:
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Similar to tRNA Isopentenylpyrophosphate transferase.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN 11) SEQUENCE FROM N.A.
 RC TISSUE=UTERUS;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC010741; AAH10741.1;
 DR InterPro: IPR002627; IPTT;
 DR InterPro: IPR000822; Znf_C2H2.

DR Pfam: PF01715; IPTT: 1.
 DR ProDom: PD004674; IPTT: 1.
 DR SMART: SM00355; Znf_C2H2: 1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 DR TRANSFERASE.
 SO SEQUENCE 324 AA; 37223 MW; 1E6835D7C09126A9 CRC64;
 Alignment Scores:
 Pred. No.: 1,11e-119 Length: 324
 Score: 1531.00 Matches: 296
 Percent Similarity: 91.10% Conservative: 1
 Best Local Similarity: 90.80% Mismatches: 1
 Query Match: 42.83% Indels: 28
 DB: 4 Gaps: 2
 US-09-513-151-3 (1-2041) x 096FJ3 (1-324)
 OY 434 ATGGGACATGAGAAAGTATGATGACCGAAAGATGAGCTTGAAGAGAGATGCTTGA 493
 DB 1 MetGlyThrGluLysValIleAspArgLysValGluLeuGluLysGluAspGlyLeuVal 20
 OY 494 CTTCAACAACGCTTAAGCCAGGTGAGCCAGAAATGCTGCCAAGTGCATCCACATGAC 553
 DB 21 LeuHisLysArgLeuSerGlnValAspProGluMetAlaIleLysLeuHisProHisAsp 40
 OY 554 AAAGCGAAATGGCCAGAGACTTGCAGATTTTGAAGAAACAGAAATCTCTCATGTA 613
 DB 41 LysArgLysValAlaArgSerLeuGlnValPheGluIleThrGlyIleSerHisSerGlu 60
 OY 614 TTTTCCATGATGCAATACAGGAAGAGTGTGTGCTCCCTGGAGGCTCTGGAAGTTC 673
 DB 61 PheLeuHisArgGlnHisThrGlnGluGlyGlyGlyProLeuGlyLysProLeuLysPhe 80
 OY 674 TTAACCCCTTCATCCCTTGGCTTCATGTCGACAGGACGCTTCAGATGACGCTTCGAT 733
 DB 81 SerAsnProCysIleLeuTrpLeuHisAlaAspIleAlaValLeuAspGluArgLeuAsp 98
 OY 734 AAGAGGCTGATGACATGCTTGGCTGCTGGGCTTCGAGAGAAAGTATGATGACGCTTCAGA 793
 DB 99 LysArgValAspAspMetLeuAlaIleGlyLeuLeuGluIleuAlaGAspPheHisArg 118
 OY 794 CGCTATATATACAGAAAGATGTTTCGGAATAATAGCCAGACTATACATGCTATCTTCAA 853
 DB 119 ArgTyrAsnGlnLysAsnValSerGluAsnSerGlnAspTyrGlnHisGlyIlePheGln 138
 OY 854 TCAATGCTTCAAGGAATTTCAAGACTACCTGATCACTGAGGAAATGACACTGAGAC 913
 DB 139 SerIleGlyPheLysGluPheHisGlyTyrLeuIleThrGlnGlyLysCysThrLeuGlu 158
 OY 914 ACTAGTAAACGCTTCTTAAGAAAGCA-----CCTGCTCCATTGTC 940
 DB 159 ThrSerAsnGlnLeuLeuLysGlyIleGluAlaLeuLysGlnValThrLysArgTyr 178
 OY 941 -----CCTGCTCCATTGTC 955
 DB 179 AlaArgLysGlnAsnArgTyrValLysAsnArgPheLeuSerArgProGlyProIleVal 198
 OY 956 CCCCCTGTATAGGCTTAGAGATGATGATGCTCGAAGTGGAGGAGTCTGTTCTTGA 1015
 DB 199 ProProValIlyrGlyLeuGluValSerAspValSerLysTrpGluGlnSerValLeuGlu 218
 OY 1016 CCTGCTCTTGAATCGTGAAGAGTTTCATCCAGGCGCACAAAGCTTACAGCCACTCCATA 1075
 DB 219 ProAlaLeuGluIleValGlnSerPheIleGlnGlyHisLysProThrAlaThrProIle 238
 OY 1076 AAGATGCTATACATGATGAGTGAAGACAGAGAAAGTATGATACCTGTGACTGTGAT 1135
 DB 239 LysMetProTyrAsnGlnAlaGlnAsnLysArgSerTyrHisLeuLysAspLeuLysAsp 258
 OY 1136 CGAATCATCATGATGGGATCCGCAATGGGCGCACAAAGTATGATACCTGTGATGAC 1195
 DB 259 ArgIleIleIleIleGlyAspArgGluTTPAlaIleHisIleLysSerLysSerHisLeuAsn 278

```

Oy 1196 CAACGTGAAGAAAGAGATGAGTGCCTGCTCAACACCATGAGAAAGTGCAGACT 1255
    |||||||
Db 279 GlnLeuLysArgArgArgLeuLeuSerSerAlaValAsnThrIleGlnSerInser 298
Oy 1256 GTTCCCGACACTATACAAAGACCTTAAGGAGGATGCCAGGCGAGATGACAA 1315
    |||||||
Db 299 ValSerProAspHisAsnLysGlnProLysGlnLysGlnSerProGlnAsnAspGln 318
Oy 1316 GAGCTGAATGCAGCGTT 1333
    |||||||
Db 319 GlnLeuLysCysSerVal 324

RESULT 5
Oy 09D1H5 PRELIMINARY: PRT: 326 AA.
AC 09D1H5:
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE 2310075G14RIK protein (RIKEN CDNA 2310075G14 gene).
CN 2310075G14RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RC MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Ishii Y.,
RA Arkawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aikawa K., Iwata M., Nishi K., Kiyosawa H., Kasukawa T., Salto R.,
RA Salto T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Salto R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleschmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK003556; BAB22853.1;
DR EMBL: BC019812; AAH19812.1;
DR MGI: MGI:1914216; 2310075G14RIK.
DR InterPro: IPR002627; IPT:
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF01715; IPT: 1.
DR ProDom: PD004674; IPT: 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN.1.
SQ SEQUENCE 326 AA: 37191 MW: 454367AAB70DD1F0 CRC64:

```

Alignment Scores:

```

Pred. No.: 7.63e-104 Length: 326
Score: 1342.00 Matches: 261
Percent Similarity: 83.74% Conservative: 12
Best Local Similarity: 80.06% Mismatches: 27
Query Match: 37.54% Indels: 26
DB: 11 Gaps: 1

```

US-09-513-151-3 (1-2041) x 09D1H5 (1-326)

```

Oy 434 ATGGCAGTGAAGAAAGTATTGATTCAGCCGAAAGTGGAGCTTGAAGAGGATGCTTGTGA 493
    |||||||
Db 1 MetGlyThrGlyLysValValAlaAspArgLysValGlnLeuGlnLysGlnAspLysGln 20
Oy 494 CTTCAAAAGCGCTAAGCCAGAGTGGAGCCAGAAATGGCTGCCAAGCTCCATCCAGATGAC 553
    |||||||
Db 21 LeuHisLysArgLeuSerGlnValAlaAspProGlnMetAlaValLeuHisProHisAsp 40
Oy 554 AAAGCGAAGTGGCCAGAGAGCTTGCAAGTTTGAAGAAAGCAAGTCTCATAGTGA 613
    |||||||
Db 41 LysArgLysValAlaArgSerLeuGlnValProGlnGlnThrGlyIleSerHisSerGln 60
Oy 614 TTTCTCCATCGTCACATACAGAGAGAGTGGCTGCCCTGGAGAGTCTGTGAAGTTC 673
    |||||||
Db 61 PheLeuHisArgGlnHisAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
Oy 674 TCTAACCTTGATCCTTTGGCTTGCTGATGCTGACAGGAGCTTGAAGAGCGCTTGAT 733
    |||||||
Db 81 ProAsnProCysIleLeuThrPheHisAlaAspGlnAlaValLeuAspGlnArgLeuAsp 100
Oy 734 AAGAGGCTGATGATGATGCTGCTGCTGGCTCTGGAGAGAGTAAAGATTTTACACA 793
    |||||||
Db 101 LysArgValAlaAspMetLeuAlaAlaGlyLeuLeuGlnGlnLeuArgGlyPheHisArg 120
Oy 794 CGCTATATCAGAGAAATGTTTCGGAATAATACCCAGAGATATCAATGATGTTTCCAA 853
    |||||||
Db 121 ArgTyrAsnLeuLysAsnLysSerGlnAsnSerGlnAspTyrGlnHisGlyIlePheGln 140
Oy 854 TCAATGGCTTCAAGAAATTTTCAGAGTACCTGATCACTGAGGAGAAATGCACTGGAG 913
    |||||||
Db 141 SerIleGlyPheLysGlnPheHisGlnTyrLeuThrThrGlnGlnLysCysThrProGln 160
Oy 914 ACTAGTACACCACTCTTAAAGAAAGCA----- 940
    |||||||
Db 161 ThrSerAsnGlnLeuLeuLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Oy 941 -----CCTGGTCCCATGCTC 955
    |||||||
Db 181 AlaArgLysGlnAsnArgTyrValLysAsnArgPheLeuSerArgProGlnProSerVal 200
Oy 956 CCCCCTGCTATGCTTACAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1015
    |||||||
Db 201 ProProValTyrGlyLeuGlnValSerAspValSerLysTyrGlnGlnSerValLeuGln 220
Oy 1016 CCGTGGCTTGAATGCTGCAAGTTTCACTGAGGCGGCAAGCAAGCTTCAAGCCATCA 1075
    |||||||
Db 221 ProAlaLeuAsnIleValGlnSerPheIleGlnGlnLysLysProThrAlaMetProVal 240
Oy 1076 AAGATGCGATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1135
    |||||||
Db 241 LysMetAlaTyrAsnGlnSerGlnLysGlnLysArgSerTyrHisMetCysAspLeuGlnAsp 260
Oy 1136 CGAATCATCATGTTGGGATGCGAATGCGAGCGGCAATGATTAATCCAAATCCATGAC 1195
    |||||||
Db 261 ArgIleIleIleGlyAspArgGlnTyrAlaAlaHisLeuLysSerLysSerHisLeuHis 280
Oy 1196 CAACGTGAAGAAAGAGATGTTGACCTGATGCTGTCGTAACACATGAAAGATGACAGT 1255
    |||||||
Db 281 GlnLeuLysLysArgTyrArgLeuSerLeuAspAlaValSerAlaThrGlySerInser 300
Oy 1256 GTTCCCGACACTATACAAAGACCTTAAGGAGGATGCCAGGCGAGATGACAA 1315
    |||||||
Db 301 AsnSerProAspCysAspProGlnArgIleGlnGlnGlnSerGlnHisAsnGln 320
Oy 1316 GAGCTGAATGCAGCGTT 1333
    |||||||
Db 321 GlnLeuLysAlaSerVal 326

RESULT 6
Oy 09GYG3 PRELIMINARY: PRT: 430 AA.
AC 09GYG3: 0950F6;

```

DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Hypothetical 49.5 kDa protein (TRNA isopentenyl transferase).
 GN ZC395.6 OR GRO-1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditidae.
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN (1)
 RN SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None.
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium."
 RL Science 282:2012-2018(1998).
 RN (2)
 RN SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Connell M.;
 RT "The sequence of C. elegans cosmid ZC395."
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RN SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN (4)
 RN SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=21444833; PubMed=11560893;
 RA Lemieux J., Lakowski B., Webb A., Meng Y., Ubach A., Bussiere F.,
 RA Barnes T., Hekimi S.;
 RT "Regulation of Physiological Rates in Caenorhabditis elegans by a
 RT tRNA-Modifying Enzyme in the Mitochondria."
 RL Genetics 159:147-157(2001).
 DR EMBL: U13642; AAG00042.2;
 DR EMBL: AY022773; AAL14112.1;
 DR InterPro: IPR002627; IPR.
 DR InterPro: IPR000822; Znf_C2H2.
 DR Pfam: PF01715; IPR1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 DR Hypothetical protein; DNA-binding; transferase; zinc-finger.
 KW SEQUENCE 430 AA: 49548 MW: 107E95095C81A2B9 CRC64;

 Alignment Scores:
 Pred. No.: 2,339-37 Length: 430
 Score: 549.00 Matches: 136
 Percent Similarity: 52.38% Conservative: 84
 Best Local Similarity: 32.38% Mismatches: 150
 Query Match: 15.36% Indels: 50
 DB: 5 Gaps: 13

 US-09-513-151-3 (1-2041) x Q9GTG3 (1-430)

 QY 71 CGGACCTACCTCTTGTAGTATTCGCGGCGACCGGCAATTCAGCGTGGC 130
 Db ||||| ||||| : : : : : ||||| ||||| ||||| ||||| |||||
 16 ArgthrasprolletlelPheValIleGlyCysthrGlythrGlyLysSerAspLeuGly 35
 QY 131 TTGCACTAGGCGGCGGCGGCGGAGATGCTGACGCGCTACTCATGACAGCTTAT 190
 Db :
 36 ValAlaIleAlaLysLysTyrGlyGlyGlyValIleSerValAspSerMetGlnPheTyr 55
 QY 191 GAAAGCTTAGACATCAACACCAAGTTTTCGCCCAAGGAGAGAAATCTGCGGCGAC 250
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 56 LysGlyLeuAspIleAlaThrAsnLysIleThrGluGluGluSerGlyGlyIleGlnHis 75
 QY 251 CACATGATCACTTTGTGATCTCTT---GTGACCAATTACACAGTGTGAGCTTCAGA 307
 Db ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
 76 HistMetSerPheLeuAsnProSerGluSerSerLysIleAsnValHisSerPheArg 95

QY 308 AATAGCAACTGCTCTGATTAAGATATATTGGCCCGGACCAAAATTCATTGTGTG 367
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 96 GluValThrLeuAspLeuIleLysIleAlaGlnAlaArgSerLysIleProValIleVal 115
 QY 368 GAGAGAACCAATTATTCATTAATCTGCTGTGGAAA-----CTTCTGTCAAT 418
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 116 GlyGlyThrThrTyrAlaGluSerValLeuTyrGluAsnAsnLeuIleGluThrAsn 135
 QY 419 ACCAAGCCCCAGAGATGGGCGACTGAGAAAGATGATGACCGAAAGTGGAGCTTGAAG 478
 Db ||||| :
 136 ThrSerAspAspValAspSerLysSerArgThrSerSerGluSerSerLysPthr 155
 QY 479 GAGGATGCTCT-----GTACTCAAAACCCCTGACGAGCGTGGACCAAGATG 529
 Db ||||| ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
 156 GluGluGlyIleSerAsnGlnGluLeuThrPaspGluLeuLysLysIleAspGluLysSer 175
 QY 530 GCTGCAACCTGATCCATCAGATGACAAACCAAGTGGCCGAGAGCTTGGCAAGTTTGA 589
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 176 AlaLeuLeuHisProAsnAsnArgTyrArgValGlnArgAlaLeuGlnIlePheArg 195
 QY 590 GAAACAGAAATCTCTATGATGATTTCCATGCTCAACATACGACAAAGAGTGTGCT 649
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 196 GluThrGlyIleArgLysSerGluLeuValGluLysGluLysSerAspGlu---ThrVal 214
 QY 650 CCCCTGGAGGCTCTGAGTCTCTTAACCCCTGATCCTTGGCTCATGCTGACAG 709
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 215 AspleGlyGlyArgLeuThrPheAspAsnSerLeuValIlePheMetAspAlaThrPro 234
 QY 710 GCAGTCTAGTACGAGCGCTTGGATAGAGGAGTGGATGACATGCTTCTGCGCTCTTG 769
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 235 GluValLeuGlnGluThrLeuAspGlyArgValAspLysMetLysLeuGlyLeuLys 254
 QY 770 GAGCACTAAGAGATTTTACACAGCGCTATATACAGAAAGATTTTGGAAATACCGAG 829
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 255 AsnGluLeuIleGluPhe-----TyrAsnGlnHis-----AlaGluTyrIleAsn 269
 QY 830 GACTATCAACAGTGTCTCTTCCATCAATTCGCTTGAAGATTCACAGAGTACCTG--- 886
 Db : : : : : : : : : : : ||||| ||||| ||||| ||||| |||||
 270 HisSerLysTyrGlyValMetGlnCysIleGlyLeuLysGluPheValProThrLeuAsn 289
 QY 887 -----ATCAGTGGAGGAAATATGC--- 904
 Db :
 290 LeuAspProSerGluArgAspThrLeuAsnGlyAspLysLeuPheLysGlnGlyCysAsp 309
 QY 905 -----ACACTGGAGACT-----ACTAACAG 925
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 310 AspValLysLeuHisThrArgGlnTyrAlaArgArgGlnArgThrPyrArgSerArg 329
 QY 926 CTCTTAAGAAAGAGACTGCTGCC---ATTGTCCTCCCTGATGCTTACGCTATCT 982
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 330 LeuLeuLysArgSerAspGlyAspArgLysMetAlaSerThrLysMetLeuAspThrSer 349
 QY 983 GATGCTCGAAGTGGAGGAGAGTCTGTTTGAACCTGCTTGAATTCGCAAGCTTTC 1042
 Db ||||| :
 350 Asp-----LysTyrArgIleIleSerAspGlyMetAspLysIleValAspGlnThr 365
 QY 1043 ATCCAGGC-----CACAAGCTACAGCCATTCATTAAGATCCATCATGAA 1093
 Db : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
 366 MetAsnGlyIleAspLeuPheGluAspIleSerThrAspThrAsn---ProIleLeuLys 384
 QY 1094 GCTGACACAAAGAGATTATCACTGCTGAGCTTGGATGAGTATCATGATGAGTGGAT 1153
 Db : : : : : : : : : : : ||||| ||||| ||||| ||||| ||||| |||||
 385 GlySerAspAlaAsnIleLeuLeuAsnGlyGluIleCysAsnIleSerMetThrGlyLys 404
 QY 1154 CGGAAATGGCGACGACATAAATCAATCCCACTTGAACCACTGAGAAAGAAAGAGA 1213
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 405 AspAsnThrProLysHisIleAspGlyLysLysHisLysHisLysHisLysGlnLysLys 424

 RESULT 7
 Q9UT75 PRELIMINARY: PRT: 434 AA.
 AC Q9UT75:
 DT 01-MAY-2000 (TREMblrel. 13, Created)

DT	01-MAY-2000 (TReMBLrel. 13, last sequence update)
DT	01-JUN-2002 (TReMBLrel. 21, last annotation update)
DE	tRNA isopentenyltransferase.
GN	SPAC343.15.
OS	Schizosaccharomyces pombe (fission yeast).
OC	Eukaryota: Fungi: Ascomycota: Schizosaccharomycetes;
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;
OC	Schizosaccharomycetes.
NC	NCBI_taxid=4896;
FN	(1)
RP	SEQUENCE FROM N.A.
RC	STRAIN=972H-;
RA	Murphy L., Harris D., Wood V., Rajandream M.A., Barrell B.G.;
RL	Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AL109739; CAB52278.1;
DR	InterPro; IPR002627; IPR1*
DR	Pfam; PF01715; IPR1;
DR	Prodom; PD004674; IPR1;
DR	TIGRfams; TIGR00174; mlaa; 1.
KW	Transferase.
SO	SEQUENCE 434 AA; 50121 MW; FCFD5101DF95F34D CRC64;
US-09-513-151-3 (1-2041) x 09UT75 (1-434)	
OY	80 CCTCTTGTAGTATCTTCGGGGCCAGGCGACCGGCAATTCACGCTGGCTTGCACSTA 139
Db	4 ProleucysValIleGlyThrIleGlyAlaGlySerLeuAlaValGlnLeu 23
OY	140 GCCACGGCGCTCGGGCGGAGATGCTACAGCGCTACATCGAGGCTGTGAAGCGCTA 195
Db	24 AlaIysArgPheGlySerGlnValIleAsnIleAspMetGlnIleTyrArgGlyPhe 43
OY	200 GACATCATCACCACCAAGGTTCTGCCCAAGACAGACGAGAAATCTCCGCGACCATCATC 259
Db	44 AspThrIleThrAsnIlyIleThrValGlnGlnIleAsnValIleHisIleArgIleMet 63
OY	260 AGCTTGTGATCTCTTGTGCACCAATTAACAATGCTGCACTTCAGAAATAGACCACT 319
Db	64 SerPheIleuAsn---PheAspIlyGlnIlySerValProGlnPheGlnIlyArgAspAlaSer 82
OY	320 GCTGTGATTAAGATATATTTGGCCCGGACAAATAATCTCTATTTGTTGGAGGACCAAT 379
Db	83 ArgValIleAspGlnIleHisSerGlnGlyIleSerIleProIleValIleGlyIlyThrHis 102
OY	380 TATTAACATTAAGTATCTGCTGGTGAAG-----GTTCTGTGC 415
Db	103 TyrTyrIleuGlnSerLeuIlePheGlnIlyAspThrThrIleuSerAlaIleAspIlyLeuThr 122
OY	416 AAT-----ACCAAGCCCGACAGAGATGGGACATGAGAAAGTATGACCGGAAA 463
Db	123 AsnAspSerSerProSerIlyArgProIleHisProAspSerHis---IleLeuAsp----- 139
OY	464 GTGAGAGCTGAAGAAGAGATGCTTTGACTTCACAAAGCGCTAAAGCCAGGTGGAGCCSA 523
Db	140 -----AspAspProSerAlaMetLeuSerTyrLeuIlyIleAspPro 154
OY	524 GAATGGCTGCCACGCGCATCCACATGACAAACGCAAAATGCGCCAGAGACTTGCAGATT 583
Db	155 ValMetIleArgIleGlnTyrPheIleAspIleArgAspThrIleArgIlyIleArgIlySerLeuGlnIle 174
OY	584 TTCTAAGAAACAGCAATCTGCATAGTGAATTTCTCCATCGTCAACATACAGGAAGAGGT 643
Db	175 TyrThrHisThrIleGlyArgProIleGlnIleIleTyrGlnIleGlnIlyMetIlySerSer 194
OY	644 GGTGGTCCCTTGGAGGCTCTGAACTTCTAAACCTTCATCTCTTGGCTTCACT 703

[illegible]

RN [3] SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Town C.D., Kaul S.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Golovko A., Hjaln G.;
 RT "A tRNA isopentenyl transferase from Arabidopsis thaliana."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Takel K., Sakakibara H., Sugiyama T.;
 RT Identification of Genes Encoding Adenylate Isopentenyltransferase, a
 RT Cytokinin Biosynthesis Enzyme, in Arabidopsis thaliana."
 RL J. Biol. Chem. 0:0-0(2001).
 DR EMBL: AC005824; AAC73024.2;
 DR EMBL: AF109376; AAF00582.1;
 DR EMBL: AB062609; BAB59042.1;
 DR InterPro: IPR002627; IPT.
 DR InterPro: IPR000822; Znf_C2H2.
 DR Pfam: PF01715; IPT; 1.
 DR ProDom: PD004674; IPT; 1.
 DR SMART: SM00355; ZnF_C2H2; 1.
 DR TIGRfams: TIGR00174; miaa; 1.
 KW transferase.
 SQ SEQUENCE 466 AA: 53098 MW: 59554DEB15FBADA CRC64;

 Alignment Scores:
 Pred. No.: 1,17e-32 Length: 466
 Score: 493.00 Matches: 133
 Percent Similarity: 50.10% Conservative: 108
 Best Local Similarity: 27.65% Mismatches: 138
 Query Match: 13.79% Indels: 102
 DB: 10 Gaps: 18

 US-09-513-151-3 (1-2041) x Q9ZUX7 (1-466)
 QY 41 CTTGTGGCAGTGGGCTCAGGGC-----CTGCACGACCCCTACTCTTGTAGTAT 94
 Db 6 CTTGTGGCAGTGGGCTCAGGGC-----CTGCACGACCCCTACTCTTGTAGTAT 94
 QY 95 CTTGTGGCAGTGGGCTCAGGGC-----CTGCACGACCCCTACTCTTGTAGTAT 94
 Db 26 Met1yProthrglyserglySerlyLeuAlaValasPheHisSer 45
 QY 155 GGTGAGATCGTACGCGCTGCTCCATGCTATGAGGCGCTGAGATCATCAGCAAC 214
 Db 46 ValGluLeuLeuAlaValasPheHisSerlySerlyLeuAlaValasPheHis 65
 QY 215 AAGTTTCTGCCAAGACGAGAACTGCGCGCACCATGATCAGCTTGTGATCT 274
 Db 66 LysValThrValasPheHisSerlySerlyLeuAlaValasPheHisSer 85
 QY 275 CTTGTGACATTTACAGAGTGGTGGTCTGAGAAATAGACCACTGCTGTGATGAGAT 334
 Db 86 Aspmet---GluPheThrAlaArgAspPheArgAspPheThrValProLeuIleGlu 104
 QY 335 ATATTCCCGGACCAAAATTCCTATGTTGTGGGAGGAGCAATTTATCATGATGATCT 394
 Db 105 IleValSerArgAsnHisIleProValLeuValGlyGlyThrHisIleGlyAla 124
 QY 395 CTGCTCTGGAAGTTCTTGTCAATACCAAGCCGAGAGATGGCACTGAGAA----- 448
 Db 125 ValValSerlyPheLeuLeuAspAlaIleGluAsp-----ThrGluGlyCys 142
 QY 449 -----GTGATTGACCGAAAGTGGAGCTGAA-----AAGGAG 481
 Db 143 AlaAspValAlaSerValAlaAspGlnAspMetValAlaGluSerValPheGlyArg 162
 QY 482 GATGCTCTTGTACTTCAAAACGCGCTAAGCAGGTGAGCCAGAAATGGCTGCAACGCTG 541

Db 163 AspLeuSerHisGlyTyrGluLeuLeuValGluLeuAspProValAlaIleAsnArg 182
 QY 542 CATCACATGACAAACGCAAAAGTGGCCAGAGCTTGCAGATTTTACAGAAACGAAATC 601
 Db 183 HisProAsnHisIleArgIleAsnGlnTyrLeuSerLeuHisIleSerArgGlyVal 202
 QY 602 TCTCATGACGAATTTCTCCATGCTCAACATACGGAAGAGGTGGCTCCCTGGAGGT 661
 Db 203 LeuProSerIlys---LeuTyrGlnGlyIleValGluAsnThrPheCysIleAsnAla 221
 QY 662 CCTGTGAGTCTCTTAACCTTGCATCCTTGGCTTCATGCTGACGAGCAGCTTACAT 721
 Db 222 Ser---ArgPhe---AspTyrCysLeuIleCysMetAlaGluThrAlaValLeuAsp 239
 QY 722 GAGCGCTTGGATAGAGGCTGATGACATGCTTGTGCTGGCTTGGAGACATAAG 781
 Db 240 ArgTyrValGluGlnArgValAlaAspAlaMetValAlaAspGlyLeuLeuAspGluVal 259
 QY 782 GATTTTCACAGACGCTATATACAGAAATGTTTCGGAATAATGCGACGCTTCAACAT 841
 Db 260 AspIleTyrIlys-----ProGlyAlaAspTyrThrArg 270
 QY 842 GGTATCTTCCATCAATATGAGCTTCAAGAAATTTACAGAGTACCTG-----ATCAGT 892
 Db 271 GlyLeuArgIleSerIleGlyValArgGluPheGluAspPheLeuValIleHisLeuSer 290
 QY 893 GAG-----GGAATATGACA---CTGCAGACTATGATACCACTTCTTAAGAAAGCA 940
 Db 291 GluThrCysAlaGlyHisLeuThrSerLeuSerAsnAspIleValMetLysGluAsn 310
 QY 941 CTTGTGCCCTT----- 952
 Db 311 LeuArgLysIleLeuAsnPheProLysAspAspLysLeuArgIleMetLeuGluAla 330
 QY 952 ----- 952
 Db 331 IleAspArgValLysLeuAsnThrArgArgLeuLeuArgGlnLysArgValSer 350
 QY 953 ---GTCCCCCTGTCTATGCTTACAGATTCGATGTC----- 968
 Db 351 ArgLeuGlnThrValPheGlyTyrAsnIleHisTyrIleAspAlaThrGluTyrIleLeu 370
 QY 989 -----TCGAAGTGGAGAGTCTGTTCTTGAACCTCTGTAACCTGTAACCTGTAAC 1036
 Db 371 SerLysSerGluGluSerThrPheAsnAlaGlnValValLysProAlaSerGluIleIleArg 390
 QY 1037 AGTTTCATCCAG-----GCCACAAAGCTTACAGCACTCCAAATTAAGATG 1081
 Db 391 CysPheLeuGlnThrGluThrGluSerGlyArgAspProThrSerGlyLys----- 407
 QY 1082 CCATACATGACGTGAGAAACAGAAAGTTATCAGCTGTGTGATCTGT---GATCGA 1138
 Db 408 -----SerIleGluArgAspLeuTyrThrGlnTyrValCysIleValCysGlyAsnLys 425
 QY 1139 ATCATGATGAGTGGCGAGTGGCGAGCGCCATCAATATCCAAATCCCACTG----- 1192
 Db 426 IleLeuArgGlyArgHisGluThrGlnHisIleLysIleGlnArgThrHisArg 445
 QY 1193 -----AACCAACTGAAAGAAAGAAAGAAAGATTCAGTACGATGCTGTC 1234
 Db 446 ThrThrArgHisLysAsnSerGlnThrTyrLysAsnArgGluValGlnGluAlaGluVal 465
 QY 1235 AAC 1237
 Db 466 Asn 466

 RESULT 9
 ID 08S926 PRELIMINARY; PRT; 439 AA.
 AC 08S926;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Similar to tRNA isopentenyltransferase.
 GN OJ1656.A11.14.
 OS *Oryza sativa* (Japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; *Oryza*.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N. A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.:
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, BAC
 RT clone:OJ1656-A11."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP003448; BAB5325.1;
 KW Transferase.
 SO SEQUENCE 439 AA: 49646 MW: 74BE689EAA2EEF95 CRC64.

Alignment Scores:
 Pred. No.: 8.75e-32 Length: 439
 Score: 482.50 Matches: 138
 Percent Similarity: 43.97% Conservative: 77
 Best Local Similarity: 28.22% Mismatches: 143
 Query Match: 13.50% Indels: 131
 Gaps: 17

US-09-513-151-3 (1-2041) x Q8S926 (1-439)

OY 11 ATGGCGTCGTCGGCGTCGACGAGCA----- 37
 ||||| :|||||: |||
 DB 1 MetAlaHisLeuAlaHisSerAlaAlaProLeuProSerAlaAspProAspAlaGlyIu 20
 OY 38 -----CTTCCTGTGGCAGTGGCGCTCAGGGCGCTCGAAGGACCCTA 79
 ||| |||||
 DB 21 GluSerSerHisSerProProProProGluLysGlyLeuArg----- 34
 OY 80 CCTCTTGATGATTCCTCGGCGCACGGGACCGCAATTCACGCTGCGTTCGAGTA 139
 :|||||:|||||: ||||| |||||: |||
 DB 35 LysValIValValIValMetCylValIthrGlyIValGlyLysSerArgLeuAlaValAspLeu 54
 OY 140 GGCCAGCGGCTCGCGGT--GAGATCGTCAGCGCTGATCCATGACAGCTATGAAGGC 196
 ||| ||| :|||||:|||||: ||||| |||
 DB 55 AlaSerHisPheAlaGlyValGluValValSerAlaAspSerMetGluValIArgIly 74
 OY 197 CTAGACATCATCACCAACAAGSTTTCTGCCCAAGACAGACAGATCTGCCGACACATG 256
 |||||:|||||: ||||| |||||: |||||
 DB 75 LeuAspValLeuThrAspHisValProLeuHisGluGlnLysGlyValProHisHisLeu 94
 OY 257 ATCAGCTTTGTGATCCCTCTGTGACCAATTCACAGCTGAGCTGAGATGAGACA 316
 :|||||:|||||: ||| :|||||: |||
 DB 95 LeuSerValIleAspProSerVal---GluPheThrCysArgAspPheArgAspHisAla 113
 OY 317 ACTGCTGTGATTAAGATATATTGCCCCGAGCAAAATTCCTATTGTTGTGGAGAGACC 376
 :|| :|| :|| :|| :||
 DB 114 ValProAlaLeuValSerProPheLeuPheAspMetAlaGlnAspIleGluIy--- 132
 OY 377 AATATTATCATGAATCTCTGCTCTGGAAGATTCTTTCATATCAAGCCCAAGAGATG 436
 :|| :|| :|| :|| :||
 DB 133 -----LeuThrIleAsnAspHisLeuAspGluIle 142
 OY 437 GGCACGTAGAAAGTATTGACCGAAAGTGAGCTTGAAGAGAGATGCTTGTGACT 496
 ||| ||| :|||||: |||
 DB 143 Gly-----LeuAspAsnAspAspIleValIleGlyLeu 152
 OY 497 CACAAACGCTAAGCCAGGTGAGCCAGAAATGCTGCCACCTGCATCATACATGACAAA 556
 :|||||: ||| :|||||: ||| :|||||: |||
 DB 153 TyrGlnHisLeuLysIleAspProValAlaIleAlaGlnArgIleHisProAsnHis 172
 OY 557 CGCAAGTGGCCAGAGCTTTCGACAGTATTTGAAGAAACAGAGATCTTCATATGTAATT 616
 |||||: ||| :|||||: ||| ||||| :|||
 DB 173 ArgLysIleLysArgTyrLeuGluLeuLeuTyrGluSerThrGlyAlaLeuProSerAspLeu 192

OY 617 CTCCATCTCAACATACGAAAGAGTGTCGTCCTTGGAGCTCTCTGAAGTTCTCT 676
 ||| |||||: ||| |||
 DB 193 PheGlnGlyAlaIleThrGluIlySerGlyArgPro-----Ser 205
 OY 677 AACCT-----TGCATCTTGGCTTCATCTGTCGACGACGACGCTTGTATGAG 724
 ||| ||| ||||| ||||| |||||
 DB 206 AsnSerArgPheAspCysCysPheLeuPheAspAlaAspLeuHisValLeuAspArg 225
 OY 725 CGCTTGATTAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 784
 :|||||:|||||: ||||| :|||||: |||||
 DB 226 TyrValAsnGluArgValAspCysMetIleAspArgGlyLeuLeuAspGluVal----- 243
 OY 785 TTTCACAGACGCTATTAATCAGAAGATTTTCGAAATTAACGACGACGATATCAATG 844
 :||| :||| :||| :||| :|||
 DB 244 -----CysAsnIleTyrAspArgGluAlaIleThrGlyGlnIly 256
 OY 845 ATCTTCAATCATATGCTGTCGACGATTTTCACGATACCTG----- 886
 :|| :|| :|| :|| :||
 DB 257 LeuArgGlnAlaIleGlyValArgIlePheAspArgIlePheArgPheArgPheAlaArg 276
 OY 887 -----ATCAGTACGAGAAATGACACA----- 907
 ||| |||||
 DB 277 LysGluThrGlyGluIleLysMetAspSerCysThrThrMetAlaGlyLeuHisAspAsp 296
 OY 908 -----CTGAGACGTAGTAACGACGCTTGAAGAAAGACGCTGCCAT 952
 :|||||: ||| :||| :||| :|||
 DB 297 AsnLeuLysGlyLeuLeuAspGluAlaValSerGlnLeuLysAlaAsnThrArgArgLeu 316
 OY 953 GTC----- 955
 |||
 DB 317 ValArgArgGlnArgArgArgLeuHisArgLeuAsnLysTyrPheGluThrAsnLeuArg 336
 OY 956 -----CCCCGTGTATGAGCTTAGAGATATGATGTCTCGAAGTGGAG 1000
 ||||| |||
 DB 337 HisIleAspAlaIleThrGluAlaIleArgIleGly-----AlaThrAlaAspSerThrAsn 353
 OY 1001 GAGTCTGTTTGAACCTGCTCTGTAATTCGTAACGTTTCATC----- 1045
 :|||||: ||| :|||||: ||||| :|||||
 DB 354 MetLysValIValLysProCysValAspIleValArgAspPheLeuSerAspArgIle 373
 OY 1046 -----CAGGCGCACAGCTACAGCCGACCTCCCAATTAATGCTACATGAA 1093
 ||| ||| :|||||: ||| |||
 DB 374 LeuAlaSerArgAspGlySerSerValThrGlySerPro--ArgMetSerSerArgIly 392
 OY 1094 GCTGAGAACAGAGAAAGTATGACCTGTCGATCTGATGAT--CGATCTCATCTTGG 1150
 :|| :|| :|| :|| :||
 DB 393 LeuThrThrGln-----TyrValCysGluAlaCysAspAsnArgValLeuArgIly 405
 OY 1151 GATCGCAATGGCGCAGCGCACATAAATCCAAATCCCACTTGACCAACTGAGAAAGA 1210
 :|||||: ||| :||| :||| :|||
 DB 410 ThrHisGluThrGluGlnHisLysGlnGlyArgCysHisArgLysArgValGlnArgLeu 429
 OY 1211 AGAAGATTGACTCAGATGCTGTCAAC 1237
 :||| :||| :||| :||| :|||
 DB 430 LysGlnLysAlaSerThrValIleSer 438

RESULT 10
 Q8R5S5 PRELIMINARY: PRT: 315 AA.
 AC Q8R5S5;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE tRNA delta(2)-isopentenyl pyrophosphate transferase.
 GN MIA OR TTE1359.
 OS Thermomacrobacter tengcongensis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Thermomacrobacteriales; Thermomacrobacteriaceae; Thermomacrobacter.
 OX NCBI_TaxID=119072;
 RN [1]
 RP SEQUENCE FROM N. A.
 RC STRAIN=MB4T / JCM11007;
 RX MEDLINE=21992816; PubMed=11997336;


```

Db 229 -----AlaAspProIleAspAlaIaIaProArg 237
OY 542 CATCCACATGACAAACGCAAGTGGCCAGACCTTGCAAGTTTGGACAAACAGCAATC 601
Db 238 AspPro-----PheAlaAspAlaAspVal 245
OY 602 TCTCAGTACGATTTCTCCATCGTCACATACGAAAGAGTGGTCCCTCGCAGCT 661
Db 246 GlyTyrArgPro----- 249
OY 662 CCTCGAAGTCTCTCAACCCCTTGATCCCTTGGCTTCATGTCAGCAGCAGCTTCAGAT 721
Db 250 AlaLeuArgPhePro---CysCysLeuLeuTrpValAspValAspAspValLeuAsp 268
OY 722 GAGCCCTGCGATAGAGGCTGATGACATGCTTGGCTGCTGGCCTTGAGAACTAAGA 781
Db 269 GlyTyrLeuAspArgArgValAspAspMetValGlyGlyMetValGlyGlyLeuGlu 288
OY 782 GATTTTCACAGACGCTATATCAGAAAGATGTTGGCAAAATAGCCAGACTATCAACAT 841
Db 289 GlyTyr-----PheAlaThrThrSerAlaSerGluArgAlaSer-----HisAla 303
OY 842 GGTACTTCCATCAATGATGCTTCAAGCAATTTCAAGCACTGATCAGTCACTGAGGAA 901
Db 304 GlyLeuGlyLysAlaIleGlyValProGlyLeuGlyAspTyrPhe-----AlaGlyArg 321
OY 902 TGCACACTGAGACTAGTATACAGCAGCTTCTAAG-----AAA 937
Db 322 LysSerLeuAspAlaAlaIleAspGluIleLysAlaAsnThrArgValLeuAlaIaIaArg 341
OY 938 GGAACCTGGTCCCAT-----GTCCCCCTGCTATAGCTTACAGTA----- 979
Db 342 GlnValGlyLysLeuArgArgMetAlaAspValTTPGlyTTPProIleArgArgLeuAsp 361
OY 980 -----TCTGATGCTCTGCAAG 994
Db 362 AlaThrAlaThrIleArgAlaArgLeuSerGlyAlaGlyArgAlaAlaIaIaAla 381
OY 995 TGGGAGAGCTGCTGTTTGAACCTGCTTGAATCGTGAATCGCAAGTTTCATC 1045
Db 382 TrpGluArgAspValArgGlyProGlyLeuAlaIaIaMetArgGlnPheVal 398

RESULT 12
OY9SB60 PRELIMINARY: PRT: 318 AA.
AC OY9SB60:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Putative tRNA Isopentenyltransferase (Adenylate
DE Isopentenyltransferase) (Cytokinin synthase) (EC 2.5.1.27).
GN F22K18.150 OR AT4G24650 OR AT1PT4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:
OC eucosids II: Brassicales: Brassicaceae: Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Medler H., Medler E., Wambutt R., Hohelsel J., Mewes H.W.,
RA Mayer K.F.X., Schueller C.;
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Medler H., Medler E., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.

```

```

RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Takei K., Sakakibara H., Sugiyama T.;
RT "Identification of Genes Encoding Adenylate Isopentenyltransferase, a
RT Cytokinin Biosynthesis Enzyme, in Arabidopsis thaliana.";
RL J. Biol. Chem. 0:0-0(2001).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=MASSILEWSKIIJA;
RA Matsumoto M., Kakimoto T.;
RL Submitted (May-2001) to the EMBL/Genbank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=MASSILEWSKIIJA;
RA Kakimoto T.;
RT "Arabidopsis cytokinin synthase.";
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL035356; CA22998.1; -
DR EMBL: AL161561; CAB79375.1; -
DR EMBL: AB062611; BAB59044.1; -
DR EMBL: AB061402; BAB59031.1; -
DR InterPro: IPR002627; IPTT.
DR Pfam: PF01715; IPTT; 1.
DR ProDom: PD004674; IPTT; 1.
KW Transiferase.
SQ SEQUENCE 318 AA: 36674 MW; BE8BD9F5E2E0F08 CRC64;

Alignment Scores:
Pred. No.: 1,02e-19 Length: 318
Score: 338.00 Matches: 97
Percent Similarity: 42.23% Conservative: 58
Best Local Similarity: 26.43% Mismatches: 90
Query Match: 9.45% Indels: 122
DB: 10 Gaps: 10

US-09-513-151-3 (1-2041) x OY9SB60 (1-318)
OY 83 CTGTAGATGATCTCTGGGGCCAGCGGCAACGCAATTCACGCTGGCTGACCTAGGC 142
Db 7 MetValValIleMetClyAlaThrGlySerGlyLysSerSerLeuSerValAspLeuAla 26
OY 143 CAGCGGCTCGCGGCTGAGATCGTCAGCGCTGACTCATCAGCTCATAGAGCCTAGAC 202
Db 27 LeuHisPheLysAlaGluIleIleAsnSerAspLysMetClnPheTyrAspLysLeuLys 46
OY 203 ATCATCACCAACAGAGTTTCTGCGCAAGACAGAGAAATCGCCGCGCACCATGATCAGC 262
Db 47 IleThrThrAsnGlnSerThrIleGluAspArgArgGlyValProHisHisLeuLeuGly 66
OY 263 TTGTGATCTCTTGTGACCAATTCACAGATGCTGACTTGAATAAGACCACTGCT 322
Db 67 GluLeuAsnProGluIaGlyGlyValIThrAlaIaGlnPheArgValMetAlaIaGlu 86
OY 323 CTGATGAAGATATATTTCGCCGACACAAATTCATGTGTGGAGAGCAATATAT 382
Db 87 AlaIleSerGluIleThrGlnArgLysLysLeuProIleLeuAlaGlyLysSerAsnSer 106
OY 383 TACATGATCTCTGCTGGAAGATTCTGTCAATACCAACCCAGAGATGGCGACT 442
Db 107 TyrIleHisAlaLeuLeuAlaLys----- 114
OY 443 GAGAAAGTATGACCGAAAGAGTGAAGTGAAGAGAGATGCTTGTACTTACAA 502
Db 114 ----- 114
OY 503 CGCTTAGCCAGGTGACCCAGAAATGCTGCCAGCTCATCCATGACAAAGCAAA 562
Db 115 -----SerTyrAspProGluAsn-----TyrProPheSerAspHisLys 127
OY 563 GTGGCCAGAGAGCTGCAAGTTTGTGAAGAAGACGAATCTCATAGTGAATTTCTCAT 622

```


Db 246 GlnleuylslyIleGlnArgLeuHisLysLysTrpLysMetSerMetHisArgValAsp 265
 Qy 977 GTATCGATGTC-----TCGAAGTCGGAGAGAGCTT 1006
 Db 266 AlathrGlnValAlpheLeuLysArgAsnValcIuGlnAspGluAlaTrpGluAsnLeu 285
 Qy 1007 GTTCTGAACCTGCTCTTGAATAGTTCGCAAGTTTCATCCAGGCCCAACAGCTTACAGCC 1066
 Db 286 ValAlaArgProSerGluAlaGllleValAspLysPheTrpAsnAsn----- 301
 Qy 1067 ACTCCAAATAAGATGCCATACATGAATGAAGCTGAGACAAGAGAACTTATCAGCTGTGCAC 1126
 Db 302 -----AsnGlnLeuLysAsnAspAspValcIuHis-Cys----- 312
 Qy 1127 CTCTGTGATCGATCATCATTTGGGAGATCGCAATGGCGAGCGCATTAAT 1178
 Db 313 -----LeuAlaIaSerTrpGlyGlyGlySerGlySerArgAlaHisAsn 327
 RESULT 14
 Q9LUG4 PRELIMINARY: PRT: 329 AA.
 AC Q9LUG4: 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE tRNA isopentenyl transferase-like protein (Adenylylate
 DE isopentenyltransferase) (EC 2.5.1.27).
 GN AtIPT7.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID-3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA.
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
 RL Submitted (Feb-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA.
 RX MEDLINE-20277480; PubMed-10819329;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 RT features of the regions of 4,504,864 bp covered by sixty p1 and TAC
 RT clones.";
 RL DNA Res. 7:131-135(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA.
 RA Takei K., Sakakibara H., Sugiyama T.;
 RT "Identification of Genes Encoding Adenylylate Isopentenyltransferase, a
 RT Cytochrome Biosynthesis Enzyme, in Arabidopsis thaliana.";
 RL J. Biol. Chem. 0:0-0(2001).
 DR EMBL: AB023036; BAB02782.1;
 DR EMBL: AB062613; BAB59046.1;
 DR InterPro: IPR002627; IPT.
 DR Pfam: PF01715; IPT.1.
 DR ProDom: PD004674; IPT.1.
 DR Transferase.
 SO SEQUENCE 329 AA; 36979 MW; 66AFA61D3DE49AF8 CRC64;

Alignment Scores:

Pred. NO.: 2,46e-19 Length: 329
 Score: 333.50 Matches: 99
 Percent Similarity: 38.94% Conservative: 56
 Best Local Similarity: 24.87% Mismatches: 105
 Query Match: 9.33% Gaps: 138
 DB: 10

US-09-513-151-3 (1-2041) x Q9LUG4 (1-329)

Qy 83 CTGTAGTATTCGGGGCCAGCGGACCGCAATTCACAGCTGGGCTTCAGACTAGGC 142

Db 36 ValIlePheValMetGlyAlaThrGlySerGlyLysSerArgLeuAlaIleAspLeuAla 55
 Qy 143 CAGCGCGTCCGGGTGAGATGTCAGCGGCTGACTCCATGCGAGGCTTATGAAGGCTTAGAC 202
 Db 56 ThrArgPheGlnGlyGluIleIleAsnSerAspLysIleGlnLeuTrpLysGlyLeuAsp 75
 Qy 203 ATCATCCCAACAGAGGTTTCCGCCAAGACAGAGAAATCTGCGGACACCATCATGAC 262
 Db 76 ValLeuThrAsnLysValThrProLysGlnLysArgGlyValProHisLysLeuGly 95
 Qy 263 TTTGTGATCTCTTGTGACCAATTACACAGTGTGAGCTTCAGAAATAGACAATGCT 322
 Db 96 ValPheAspSerGluAlaGlyAsnLeuThrAlaThrGlnTrpGlnTrpArgLeuAlaSerGln 115
 Qy 323 CTGATTCAGATTAATTTGCCCCGAGACAAATTCCTATTGTTGGCGAGACCAATTTAT 382
 Db 116 AlaIleSerLysLeuSerAlaAsnAsnLysLeuProIleValAlaGlyLysSerAsnSer 135
 Qy 383 TACATTAATCTCTCTGGAAGTCTTGTCAATACCAAGCCCCAGAGATGGGCACT 442
 Db 136 TyrIleGluAlaLeu----- 140
 Qy 443 GAGAAAGTATGACCGAAAGTGAGCTGAAAGAGAGATGGCTTACTTCACAAA 502
 Db 140 ----- 140
 Qy 503 CGCCTAAGCCAGGTGGACCCAGAAATGCTGCCAAGCTGCATCCACATGACAAACGCAA 562
 Db 140 ----- 140
 Qy 563 GTGGCCAGAGACTTGCAAGTTTGAAGAACAGAAATCTCATAGTGAATTTCTCAT 622
 Db 141 -----ValAsnHisSer----- 145
 Qy 623 CGTCAATACGAGAAAGGTGGTCCCTTGAGAGCTCTGAAATGTTCTTAACCT 682
 Db 146 -----GlyPheLeuLeuAsnAsnTrpAspCys 154
 Qy 683 TGCATCCTTTGGCTTCATGCTGACAGGCACTTTCATGAGCCCTGGATGAAGAGCTG 742
 Db 155 CysPheIleTrpValAspValSerLeuProValLeuAsnSerPheValSerLysArgVal 174
 Qy 743 GATACATCGCTTGGCTGGGCTCTGGAGAGCAATGAGATTTTCACAGCGCTTAAT 802
 Db 175 AspArgMetMetGluAlaGlyLeuLeuGlnIuValArgLeuAla-----PheAsn 191
 Qy 803 CAGAGAAATGTTTGGAAATAGCCAGAGCTATCAACATGATCTTCAATCAATAGGC 862
 Db 192 ProLys-----AlaAsnTrpSerValcIuIleArgGlyAlaIleGly 205
 Qy 863 TTCAAGCAATTTACAGAGTACCTGATGATCAGGGAATTCACACAGCTAGTAAC 922
 Db 206 ValProGluLeuHisGlyTrpLeuArgAsnIleSerLeuValAspArgAlaThrLysSer 225
 Qy 923 CAGCTTTTA-----AAGAAGAGACCTGGTCCCATTTGTC----- 955
 Db 226 LysMetLeuAspValAlaValLysAsnIleLysLysAsnThrGluIleLeuAlaCysArg 245
 Qy 956 -----CCCTGTCTATGCTTACAG 976
 Db 246 GlnleuLysLysIleGlnArgLeuHisLysLysTrpLysMetSerMetHisArgValAsp 265
 Qy 977 GTATCGATGTC-----TCGAAGTCGGAGAGAGCTT 1006
 Db 266 AlathrGlnValAlpheLeuLysArgAsnValcIuGlnAspGluAlaTrpGluAsnLeu 285
 Qy 1007 GTTCTGAACCTGCTCTTGAATAGTTCGCAAGTTTCATCCAGGCCCAACAGCTTACAGCC 1066
 Db 286 ValAlaArgProSerGluAlaGllleValAspLysPheTrpAsnAsn----- 301
 Qy 1067 ACTCCAAATAAGATGCCATACATGAATGAAGCTGAGACAAGAGAACTTATCAGCTGTGCAC 1126

```

Db 302 ----- AsnGlnLeuLysAsnSnpAspValGluHis-Cys----- 312
QY 1127 CTCGTGATGCATATCATTCATTGGGAGATGCGCAATGCGCAGCGCATTAAT 1178
Db 313 ----- LeuAlaIaSerTyrGlyGlyGlySerArgLahHisAsn 327
RESULT 15
Q94ID2 PRELIMINARY; PRT; 330 AA.
ID Q94ID2 AC
AC Q94ID2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Cytochrome synthase.
GN AtTPS.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RX (1)
RP SEQUENCE FROM N.A.
RC STRAIN=MASSILEMSKITA;
RA Matsumoto M., Kakimoto T.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=MASSILEMSKITA;
RA Kakimoto T.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB061403; BAB59032.1;
DR InterPro; IPR002627; IPT.
DR Pfam; PF01715; IPT.
DR ProDom; PD004674; IPT.
SO SEQUENCE 330 AA; 37408 MW; 0F440B0A4237CD02 CRC64;

Alignment Scores:
Pred. No.: 6,45e-19 Length: 330
Score: 328.50 Matches: 94
Percent Similarity: 38.38% Conservative: 53
Best local Similarity: 24.54% Mismatches: 105
Query Match: 9.19% Indels: 131
Db: 10 Gaps: 7

US-09-513-151-3 (1-2041) x Q94ID2 (1-330)
QY 20 GTGGCGGCTGCACGACGAGTT-----CTGTGGCGAGTGCGCTCAAGGCGCTGCAACGG 73
Db 5 MethrAlaLeuArgGlnValIleGlnProLeuSerLeuAsnPhgGlnCysAsnMetVal 24
QY 74 ACCCTACT-----CTGTACTGATTCCTCGGGCGGCGACGGCACC 112
Db 25 AspValProPhePheIaArgValGlySAspLysValValPheValMetGlyAlaTrnGlyThr 44
QY 113 GCGAATTCACGCGTGGCGTTGACGCTAGCCAGCGGCTCGCGGTGAGATGTCAGGCT 172
Db 45 GlySserArgLeuAlaIleAspLeuAlaThrArpPheProAlaGluIleValAsnSer 64
QY 173 GACTCCATGCAGGCTTATGAAGGCTTAGACATCATCACCAACAAGGTTTGGCCACAAG 232
Db 65 AspLysIleGlnValTyrLysGlyLeuAspIleValAlaThrAsnLysValThrProGluGlu 84
QY 233 CAGAGATTCGCGCGGACCAATGATGAGCTTTGTGATGCTCTTGACCAATTACACA 292
Db 85 SerLeuGlyValProHisIleLeuGlyTyrHisAspThrTyrGlyAspPheThr 104
QY 293 GTGTGCACTTCACAATAATAGACCAACTGCTCTGATTCAGATATATTATGGCCAGACAA 352
Db 105 AlaGluAspPheGlnArgGluAlaIleArgAlaValGluSerIleValAlaGlnArgAspArg 124
QY 353 ATTCCATTGTGTGGGAGACACCAATTATTACATTTGAATCTGTGCTTGGAAAGTTCTT 412

```

Db	125	VaIProllellellaaglygylserAsnSettyrIlegluAlaLeuVal	-----	140
Oy	413	GTCAATACCAAGCCCAAGAGATGGCACTAGAGAAAGTATGACCGAAAGTGGAGCTT		472
Db	141	-----	AsnAspGysValAspPheAtgLeuAtgTyr	150
Oy	473	GAAGAAGAGATGGCTTTGTACTTCACAAAGCCCTTAAGCCAGGTGGACCCAGAATGGCT		532
Db	150	-----		150
Oy	533	GCCAAAGTCGATCCACATGACAAAGCAAGTGGCCAGAGCTTGCAGATTTTGAAGA		592
Db	150	-----		150
Oy	593	ACAGAAATCTCATAGTAGAATTTCTCCATGTCACATACGGAAGAAGTGGTGTCC		652
Db	150	-----		150
Oy	653	CTTGAGAGTCCCTGAAGTTCTTAACCCCTTGACCTTTGGCTTACGTGCACAGCA		712
Db	151	-----	AsnCysGysPheLeuTyrValAspValSerAtgPro	162
Oy	713	GTCTAGATGAGCCGTTGGATTAAGAGGGTGGATGACATGCTTCCTGCGGCTTGGAG		772
Db	163	ValLeuHisSerPheValSerGluArgValAspLysMetValGlnMetGlyLeuValAsp		182
Oy	773	GAACATAAGAGATTTTTCACACAGCCTAATACAGAGATGTTTGGAAATACCCAGGAC		832
Db	183	GluValArg	ArgIlePheAspProSerSerSerAsp	194
Oy	833	TATCAACATGGTATCTTCCATCAATATGCGCTTCAAGAAATTTTACACAGTACCTGATCACT		892
Db	195	TyrSerAlaGlyIleAlaGluAlaIleGlyValProGlnLeuAspGluPheLeuAspSer		214
Oy	893	GAGGAGAAA--TGCACACATGACACTAGTAACACAGCTTAAAGAAAGACCTGGTCC		949
Db	215	GluMetArgAsnTyrProAlaGlnThrGlnGluArgLeuLeuGlnThrAlaIleGluLys		234
Oy	950	ATTGTCCCCCTGTCTATGGCTTGAAGAGTATGATGATCTCGAAG--		994
Db	235	IleLysGluAsnThrCysLeuLeuAlaCysArgGlnLeuGlnLysIleGlnArgLeuTyr		254
Oy	994	-----		994
Db	255	LysGlnTPrLysTPrAsnMetHisArgValAspAlaThrGlnValPheLeuArgArgGly		274
Oy	995	-----	TGGAGAGAGTCTGTCTTGAACCTGCTCTTGAATCGTGCA	1033
Db	275	GluGluAlaAspGluAlaTPrAspAsnSerValAlaHisProSerAlaLeuAlaValGln		294
Oy	1037	AGTTTCATC		1045
Db	295	LysPheLeu		297

Search completed: April 21, 2003, 18:55:40
Job time : 108.298 secs

Score: 1074.00 Matches: 214
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.17% Mismatches: 3
Query Match: 30.04% Indels: 1
DB: 10 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-764-853-586 (1-222)

QY 2 TGGCATAAGATGGCGTCCGCGGGGCTGCACGAGCAGTTCCTGTGGCAGTGGCTCAGG 61
DB 1 CysH1sLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
QY 62 GGCCTGCACAGCAGCCTTACCTTTGTAGTATCTTCGGGGCGCCAGGGCAGGCAATCC 121
DB 21 GlyLeuGlnArgThrLeuProLeuValAlaIleLeuGlyAlaThrGlyThrGlyLysSer 40
QY 122 AGCGTGGCGTTGCAGTAGCGCCAGCGGCTCGCGGTGAGATCGTCAGCGCTGATCCATG 181
DB 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGluIleValSerAlaAspSerMet 60
QY 182 CAGGCTATGAGAGCGCTTACATCATCACCACAGAGTTTCTGCCAAGACAGAGATC 241
DB 61 GlnValTyrlGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80
QY 242 TGGCGGACACATGATGATGCTTGTGATCCTCTGTGACCATATACAGAGTGGAGC 301
DB 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrlThrValAlaAsp 100
QY 302 TTCAGAAATAGACCAACTGCTGTGATTAAGATATATTTGCCGACACAAATTCCTATT 361
DB 101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle 120
QY 362 GTTGTGGGAGGAACAATTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 421
DB 121 ValValGly**ThrAsnTyrlTyrlGlnSerLeuLeuTrpLysValLeuValAsnThr 140
QY 422 AACCCGACAGAGATGGCAGTGAAGATGATGACCGAAATGGAGCTTGAAGAGAG 481
DB 141 LysProGlnGlu**GlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysGlu 160
QY 482 GATGCTCTGTACTTTCACAAGCGCTTAAGCAGGTGGAGCCAGAAATGGCTGCCAAGCTG 541
DB 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaAlaLysLeu 180
QY 542 CATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGATTGTTGAAGAACAGAGATC 601
DB 181 HisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGlnGluThrGlyIle 200
QY 602 TCTCATAGTGAATTTCTCATGCTCAACATAC-GGAAGAAGCGTGGTCCCT 654
DB 201 SerHisSerGlnPheLeuHisArgGlnHisThrGlyArgLysTrpIlePro 218

RESULT 2
US-09-764-853-784
Sequence 784, Application US/09764853
Patent No. US20020090672A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P1206
CURRENT APPLICATION NUMBER: US/09/764,853
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 939
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 784
LENGTH: 222
TYPE: PRM
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (124)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE
LOCATION: (145)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-784

Alignment Scores:
Pred. No.: 8-8e-94 Length: 222
Score: 1074.00 Matches: 214
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.17% Mismatches: 3
Query Match: 30.04% Indels: 1
DB: 10 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-764-853-784 (1-222)

QY 2 TGGCATAAGATGGCGTCCGCGGGGCTGCACGAGCAGTTCCTGTGGCAGTGGCTCAGG 61
DB 1 CysH1sLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
QY 62 GGCCTGCACAGCAGCCTTACCTTTGTAGTATCTTCGGGGCGCCAGGGCAGGCAATCC 121
DB 21 GlyLeuGlnArgThrLeuProLeuValAlaIleLeuGlyAlaThrGlyThrGlyLysSer 40
QY 122 AGCGTGGCGTTGCAGTAGCGCCAGCGGCTCGCGGTGAGATCGTCAGCGCTGATCCATG 181
DB 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGluIleValSerAlaAspSerMet 60
QY 182 CAGGCTATGAGAGCGCTTACATCATCACCACAGAGTTTCTGCCAAGACAGAGATC 241
DB 61 GlnValTyrlGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80
QY 242 TGGCGGACACATGATGATGCTTGTGATCCTCTGTGACCATATACAGAGTGGAGC 301
DB 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrlThrValAlaAsp 100
QY 302 TTCAGAAATAGACCAACTGCTGTGATTAAGATATATTTGCCGACACAAATTCCTATT 361
DB 101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle 120
QY 362 GTTGTGGGAGGAACAATTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 421
DB 121 ValValGly**ThrAsnTyrlTyrlGlnSerLeuLeuTrpLysValLeuValAsnThr 140
QY 422 AACCCGACAGAGATGGCAGTGAAGATGATGACCGAAATGGAGCTTGAAGAGAG 481
DB 141 LysProGlnGlu**GlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysGlu 160
QY 482 GATGCTCTGTACTTTCACAAGCGCTTAAGCAGGTGGAGCCAGAAATGGCTGCCAAGCTG 541
DB 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaAlaLysLeu 180
QY 542 CATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGATTGTTGAAGAACAGAGATC 601
DB 181 HisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGlnGluThrGlyIle 200
QY 602 TCTCATAGTGAATTTCTCATGCTCAACATAC-GGAAGAAGCGTGGTCCCT 654
DB 201 SerHisSerGlnPheLeuHisArgGlnHisThrGlyArgLysTrpIlePro 218

RESULT 3
US-10-260-877-32
Sequence 32, Application US/10260877
Publication No. US20030021813A1
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Chovan, Linda E.
APPLICANT: Hessler, Paul E.
APPLICANT: Reich, Karl A.
TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME
TITLE OF INVENTION: SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION O
FILE REFERENCE: 6565.US.PI
CURRENT APPLICATION NUMBER: US/10/260,877


```

Db 104 ValGlnSerLeuValAlaAspAspTTP-----GlnPheProPro 115
QY 440 ACTAGAGAAAGTGAATGCCAAAGAGTGAAGAGAT---GGCTTT---GTA 493
Db 116 ThrAspSerAlaValAlaArgPheGlnAlaArgLeuAlaAspIleGlyValGluAla 135
QY 494 CTTCCAAAGCGCTAAGCCAGGTGAGCAAAATGGCTGCAAGCTGCATCCAGATGAC 553
Db 136 LeuHisAlaGluLeuThrGlnLeuAspProGluAlaAlaAlaIleGluSerAspAsp 155
QY 554 AAACGCAAGTGGCCAGGAGCTTGCAAGTTTGGAGAAACAGGAATCTCATAGTGA 613
Db 156 ProArgArgThrValAlaArgAlaLeuGluValIleGluLeuThrGly----- 170
QY 614 TTTCTCATGCTGACATGAGGAGAGAGGTGGTCCCTTGAGAGCTCTGAGATTC 673
Db 171 -----GlnProPheGlnAlaSerGlnProProLysAspAlaProArgTTP 186
QY 674 TCTAACCTTGATCCTTTGCTTCATGTCAGCAGCAGTTCATGATGAGCGCTTGAT 733
Db 187 GltThrArg---IleIleGlyLeuLysThrThrProGluThrPheLeuAsnProArgIleGlu 205
QY 734 AAGAGGCTGATGACATGCTTCTGCTGCTGCTGCTGAGAGACTA 778
Db 206 GlnArgThrAlaArgMetPheGlnGlyPheValAlaGluVal 220

```

RESULT 5

```

US-09-864-761-42725
; Sequence 42725, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408

```

```

; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42725
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL033527.25
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN BT4/4, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: BE242161.1, EVALU 3.00e-21
; OTHER INFORMATION: SWISSPROT HIT: Q9KAC3, EVALU 2.00e-04
US-09-864-761-42725

```

Alignment Scores:

```

Pred. No.: 4,76e-15 Length: 57
Score: 246.00 Matches: 48
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.88% Indels: 0
DB: 10 Gaps: 0

```

US-09-513-151-3 (1-2041) x US-09-864-761-42725 (1-57)

```

QY 185 GCTATGAGAGCGCTAGACATCATCCACACAGCTTTGCGCCAGACAGACATCTGC 244
Db 1 ValTtGluGlyLeuAspIleThrAsnLysValSerAlaGlnGlnArgIleCys 20
QY 245 CGGACACACATGATCAGCTTGTGATCCCTTGACCAATACAGAGTGGAGCTTC 304
Db 21 ArGHisHisMetLisSerPheValAspProLeuValThrAsnThrValValAspPhe 40
QY 305 AGAATAGAGCAACTGCTGTGATT 328
Db 41 ArgAsnArgAlaThrAlaLeuIle 48

```

RESULT 6

```

US-09-881-752A-230
; Sequence 230, Application US/09881752A
; Patent No. US20020115078A1
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020115078A1el Helicobacter Polypeptides in
; FILE REFERENCE: 06132/041002
; CURRENT APPLICATION NUMBER: US/09/881,752A
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/833,457
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 230
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-881-752A-230

```

Alignment Scores:

```

Pred. No.: 3,13e-07 Length: 218
Score: 167.00 Matches: 55

```

Percent Similarity: 50.25%
 Best Local Similarity: 27.09%
 Query Match: 4.67%
 DB: 10
 Conservative: 47
 Mismatches: 61
 Indels: 40
 Gaps: 10

US-09-513-151-3 (1-2041) x US-09-881-752A-230 (1-218)

OY 350 AAAATTCCTATTGTGGAGAGAACCAATTATTCATTTGATCTCTGCTGGAAGT 409
 Db 7 GluileuLeuileValGlyserSerPheTyrluLeuLysSerileuLcLulY---- 25
 OY 410 CTTCGATACCAAGCCCCAGAGAGTGGCACTGAGAAAGTATTGACCGAAAGTGAG 469
 Db 26 --LeuSerArgMetProLysLeuSerGlyGluGlu-----ValValLys 39
 OY 470 CTGGAAGAGAGATGCTTCTTACTTCACAAAGC-----CTAAGCAGGTGACCCA 523
 Db 40 lIeGluArgGluileAlaThrLeuSerAsnProTyrllePheLeuLysSerileaspPro 59
 OY 524 GAAATGCTGCCACAGCTCATCCACATGACAAACGCAAGTGGCCAGCTTGGACATT 583
 Db 60 AsnMetAlaIlePheLysIleHisProAsnSprHrTyrlArgHnHisLysAlaLeuLcLul 79
 OY 584 TTGAAGAAGACAGGAATCTCTCATGATGATTTCTCATGTCATGCAACATACGGAAGAGT 643
 Db 80 PheTyrlAlaThrCysThrProProSerGluTyrlPheLys----- 92
 OY 644 GGTGCTCCCTTGGAGGCTCTGAG---TTCTTAACCTTGCATCTTTGG---CTT 697
 Db 93 -----AlaAsnProLysLysProPheGluHisAlaIleSerLeuPheAlaLeu 108
 OY 698 CATGCTGACGAGGAGCTTTAGATAGCGCTTGATAGAGGCTGATGACATGCTTGT 757
 Db 109 SerileGluLysSerAlaLeuHisAsnIleLysArgAlaGlyHnLysAsnMetLeuHis 128
 OY 758 GCTGGCTCTTGGAGAGCACTAAGATTTTCACAGACGCTATTAATCAGAAAGATTTTCG 817
 Db 129 SerGlyLeuValGluGluIleLysAlaLeuTyrlHnGlnTy-----Pro 143
 OY 818 GAAATACCAAGACTATACCATGTATCTTCCATCATCAATTTGCTTCAAGAAATTTTAC 877
 Db 144 LysAspSerGlnPro-----PheLysAlaIleGlyValLysGlu----- 156
 OY 878 GAGTACGATCATCTGAGGAAATGCACACTGAG-----ACTAGT 919
 Db 157 SerValLeuPheLeuGluLysArgLeuThrLeuLysGluLeuGluAlaIleThrSer 176
 OY 920 AACCAAGCTT 928
 Db 177 AsnThrMet 179

RESULT 7

US-09-759-508B-2

Sequence 2, Application US/09759508B

Publication No. US20020182599A1

GENERAL INFORMATION:

APPLICANT: Fishman, Mark C.

TITLE OF INVENTION: Methods for Diagnosing and Treating Heart Disease

FILE REFERENCE: 00786/381002

CURRENT APPLICATION NUMBER: US/09/759, 508B

PRIOR FILING DATE: 2001-01-12

PRIOR APPLICATION NUMBER: US 60/175,787

NUMBER OF SEQ ID NOS: 11

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 26926

TYPE: PRT

ORGANISM: Homo sapiens

US-09-759-508B-2

Alignment Scores:

Pred. No.: 0.232

Length:

26926

Score: 116.00
 Percent Similarity: 33.21%
 Best Local Similarity: 20.49%
 Query Match: 3.24%
 DB: 9
 Matches: 108
 Conservative: 67
 Mismatches: 183
 Indels: 169
 Gaps: 25

US-09-513-151-3 (1-2041) x US-09-759-508B-2 (1-26926)

OY 13 GCGGCGCGTGGCGGCTGC-----ACGAGCAGTTCGTTGGG 48
 Db 7981 GlyLeuProGlyProCysLysAspIleLysAlaSerAspIleHnLysSerCysLys 8000
 OY 49 CAGTGGGTCAGGGGCTGCAGACGACCTTACCTTGTAGTATTCGGGGCCAGCGG 108
 Db 8001 LeuThrTrpGluProProGluPheAsp-----GlyGly 8011
 OY 109 CACCGCAATTCACAGCTGGCTGCTCAGCTAGCCAGCGGCTGGCGGTGAG----- 160
 Db 8012 ThrProIleLeuHisTyrlValLeuGluArgGluAlaGlyArg-THrTyrllePr 8031
 OY 161 -ATGTCAGCGCTGACCTCCATGCAAGCTCATGAGCCTAGACATCATC----- 208
 Db 8031 ovalMetSerGlyGluAsnLysLeuSerTrpHnValLysAspIleProAsnGlyG1 8051
 OY 209 -----ACCAACAGGTTCTGCCCAAGACAGACAGAACTGCCG 246
 Db 8051 uTyrlPhePheArgValLysAlaValAsnLysAlaGlyGlyGlyTyrlleGluLeuLys 8071
 OY 247 GCACCAATGATCAGCTTGTGATGCTCTTGTGACCAATTACACAGTGGCTGACTTACG 306
 Db 8071 sAsnProValIleAla---GlnAspProLysGlnProProAspProproValAspValG1 8090
 OY 307 A---AATAGACCACTGCTGTGATGAGATATTTGCGCCAGCAAAATTCCTATGCT 363
 Db 8090 uValHisAsnProThrAla---GluAlaMetHnThrTrpLysProProLeuTy 8108
 OY 364 TGTGGAGAGAAC-----AATTATTACATGATCTGCTC----- 400
 Db 8108 raspolYelYserLysIleMetGlyTyrlleIleLysIleAlaLysGlyGluGluArg 8128
 OY 401 -TGAAA-----GTCTTCTCAATVACC---AAGCCCCAGA 432
 Db 8128 gTrpLysArgCysAsnGluHisLeuValProIleLeuHnTyrlHnAlaLysGlyLeuG1 8148
 OY 433 GATGGCACTAGAAAGATGATGACCGAAAGTGGAGCTTGAAGAAGAGATGCTTGT 492
 Db 8148 uGluGlyLysGlu-----TyrlGlnPheArgValArgAlaGluAsnAlaValGlyLeu 8166
 OY 493 ACTTCACAAAGCCTTACG-----CAGGTGAGCCAGCAAGATGCTGCCAGCT 540
 Db 8166 rGluProSerArgAlaThrProProThrLysAlaValAspProIleAspAlaProLysVa 8186
 OY 541 GCATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGATTTTGAAGAACAGCAAT 600
 Db 8186 lIleLeuArgHnSerLeuGluValLysArgGlyAspGluIleAlaLeuAspAlaSerIl 8206
 OY 601 CTTCATATGTAATTTCTCCATCGTCAACATACGAGAGAGGTGTGCTCCCTTGGAGG 660
 Db 8206 eSerClySerProTyrl----- 8211
 OY 661 TCCTCTGAAGTTCCTTAACCTTGCATCCTTGGCTTCATGCTAGCCAGGAGTTCTAGA 720
 Db 8212 -----ProHnThrTrpLysAspGluAsnValIleValPr 8225
 OY 721 TGAGCGCTTGATAGAGGTGATGATGCTT-----CGTGGCGCTCTTGGAGAGA 774
 Db 8225 oGluGluIleLysLysArgAlaAlaProLeuValArgArgArgLysGluValGlnG1 8245
 OY 775 ACTAAGAGATTTTCACAGACCTTATATCAACAGAAATGTTCCGAAATACCGAGACTA 834
 Db 8245 uGluGluProPheValLeuProLeuThrGlnArgLeuSerIleAspAsnSerLysLysG1 8265
 OY 835 TCAA-----CATGCTATCTTCCATCAAT 858

Db 8265 yGlsSerGlnLeuArgValArgAspSerLeuArgProAspHisGlyLeuTyrMet----- 8283
OY 859 TGCGTTAAGCAATTTCACAGTACTGATCAGTACAGGAGAAATGCACACTGAGAGACTAG 918
Db 8284 ----lIeIySaValGluAsnAspHisGlylIeAlaIySaIAProCysThrValSerValle 8302
OY 919 TAACCACCTTCTAAGAAAGAGACTGCTCCCATTTGCCCCCTCTATATGC-----TT 972
Db 8302 uAsPrThr-----ProGlyProProIleAsnPhenAlaIleGluAspIleAr 8317
OY 973 AGAGATATCTGATCTCTGCAAGTGGAGAGTCTGTTCTTGAACCTGCTCTGAATGCT 1032
Db 8317 gLyStHrSerValLeuCysIyStrpGluProProLeuAspAspGlyGlySerGluIleI 8337
OY 1033 GCAAAATTCATCCAGGAGGCAAAAGCCTACAGCCATCTCAATAAAGATGCCATCAATGA 1052
Db 8337 eAsnTyrThrLeuGlu---LysIySaSPlYsThrIySaPro----- 8349
OY 1093 AGCGAGACAGAGAGAGTTATACCTGTGTGACCTGTGTGATCATCATTTGGGGA 1152
Db 8350 -----As 8350
OY 1153 TCAGCAATGGCGACG-----CAC----- 1171
Db 8350 pSerGluTrpIleValIValThrSerThrLeuArgHisCysIyStrSerValThrIySle 8370
OY 1172 -ATAAATCCAAATCCCACTTGAACCACTGAGAGAAAGAAAGAGATTG----- 1219
Db 8370 uIleGlIyGlySglIyStrLeuPheArgValArgIleGluAsnArGpHeGlyProGlyPr 8390
OY 1220 -----GA 1221
Db 8390 oProCysValSerIySPlroLeuValAlaIySaSPlroPheGlyProProAspAlaProAs 8410
OY 1222 CTCAGATGCTGTCAACACCATAGAGTGTTCCTCCACACTATACAAGAAC 1281
Db 8410 pLySPlroIleValIleuAspValThrSerAsnSerMetLeuValIyStrIpaSn---GluPr 8429
OY 1282 TAAAGGAGAGGATCCCA 1300
Db 8429 oLySaSPlroGlySerPro 8435

RESULT 8
US-09-828-447-12
Sequence 12, Application US/09828447
Patent No. US2002069432A1
GENERAL INFORMATION:
APPLICANT: COSTA E SILVA, OSMALDO DA
APPLICANT: BOHNET, HANS J.
APPLICANT: VAN THIELEN, NOCHA
APPLICANT: CHEN, RODYING
APPLICANT: ISHITANI, MANABU
TITLE OF INVENTION: SIGNAL TRANSDUCTION STRESS-RELATED PROTEINS AND METHODS
FILE REFERENCE: 16313-0037
CURRENT APPLICATION NUMBER: US/09/828,447
CURRENT FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: 60/196,001
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 628
TYPE: PRT
ORGANISM: Physcomitrella patens
US-09-828-447-12

Alignment Scores:
Pred. No.: 0.0477 Length: 628
Score: 115.00 Matches: 101
Percent Similarity: 34.39% Conservative: 83
Best Local Similarity: 18.88% Mismatches: 169

Query Match: 3.22% Indels: 182
DB: 10 Gaps: 26
US-09-513-151-3 (1-2041) x US-09-828-447-12 (1-628)
OY 146 CGGCTCGCGGTGATGATCGTACGCGGTAC-----TCCATGCG 184
Db 143 ArgArgGlyAlaArgValAlaGluLeuAspLeuThrProAspAspIySglIyMetIyS 162
OY 185 GTCTATGAGAGCCTAGACATCATCACACAGGTTTGCCTCA----- 229
Db 163 ValThrHisGly---AsnThrLeuThrAsnProValSerPheGlnIyScysValThrAla 181
OY 230 -----GACACAGAAATCTCCCGGACCCATGATCAGC 262
Db 182 lIeIySaAsnAlaPhePheThrSerGluTyrProValCys-----ValThr 197
OY 263 TTTGTGATCTCTTGGACCAATTACACAGTGTGACTTCAGAAATAGACCACTGCT 322
Db 198 lIeGluAspHisLeuThrSer-----GluLeuGlnGlyHisAlaIleGlu 212
OY 323 CTGATTGACAGATATATTGCGCCAGACAAATTCCTATTGTTGTGGAGAGAACCAATTAT 382
Db 213 lIeLeuGlnGlnIle-----LeuGlyAspAlaLeuTyr 223
OY 383 TAC-----ATGAATCTCTGCTCGGAA 406
Db 224 TyrProProThrThrAspAlaLeuValGluPheProSerProGluSerLeuIySArgIyS 243
OY 407 GTTCTTGTCAATACCAAGCCCGAGAGATGGCAGTGAAGATGATTCACGAAAGT 466
Db 244 lIeIleIleSerThrIySPlroProIySglIyTyrLeuGlnIyLacysSerThrGlnIySleu 263
OY 467 -----GACCTGAAAGAGAGATGCTTACTT--- 496
Db 264 AlameGluAsnArgAsnLeuValAlaGluIleGluIySglIySPlroIySLeuGlnIyThr 283
OY 497 -----CACAAAGCCTTAAAGCAGGTGAGACCAGAAATAGCGTGC 535
Db 284 ThrPheAlaProLeuGlnIyLysAsnHisIleLeuGlnIyLysAsnThrProSerLeuArgIyS 303
OY 536 AAGCTGCATCCACATGACAAACGCAAGTGGCCAGAGCTTCGACATTTTGAAGAAC 595
Db 304 GluValIleuValLeuSerGlnIySglIyMetSerThrProIleGluLeuAsnSerArgSer 323
OY 596 GGAATCTCTATAGTGAATTTCTGCATCTGCATACATACGAGAACAGGTGCTGCCCTT 655
Db 324 ProSerAspLeuGlnIyAlaThrSerThrArgTyrSerIySerAsnAspGlyAsn--- 342
OY 656 GAGGCTCCTGAAAGTCTTAACCCCTGCATCTTGCGCT-----CATGCT 703
Db 343 AspAsnProIyShiSPlroIySglIyValArgLeuIleThrIleArgLeuAlaIyShiAla 362
OY 704 GACCAGGCACTTCTAGATGAGCGCTTGAT-----AAGAGGCTG 742
Db 363 LysGlyThrSerMetGluHisArgLeuGlnValAspGluSerValIySArgIleSerLeu 382
OY 743 GATGACATGCTTGTCTGCTGGCGCTTGGAGGAACTAAGACATTTTACAGACCTTAAAT 802
Db 383 SerCysLeuSerIySglIyLysValIleGluIyStrpProGluAlaLeuValIySPhenThr 402
OY 803 CACAGAGAT----- 811
Db 403 GluIySaSnIleLeuArgValTyrProAlaAlaAsnArgValAsnSerSerAsnPhenCys 422
OY 812 -----GTTTCGAAATATAGCCAGCATATCA 838
Db 423 ProThrLeuAlaTrpAsnTyrGlyAlaGlnMetValAlaGlnAsnMetGlnIyTyrGly 442
OY 839 CATGATATCTTCAATCAATTTGCG---TTCAGAGAAATTCACAGATCTGATCACTGAG 895
Db 443 LysGluLeuTrpIleAlaIlePheGlyIySPhenS-----GlyAsn 455


```

Db      495  IIESeqGlySerSercIylLeuSerAspSergIuileGu 507
          |||||  |||  ::|||  ||||||:::
RESULT  11
US-09-134-333-11
; Sequence 11, Application US/09134333
; Patent No. US20020076403A1
GENERAL INFORMATION:
APPLICANT: LONGCRE-ANDRE, SHIRLEY
APPLICANT: ROTH, CHARLES
APPLICANT: NATO, FARIDABANO
APPLICANT: BARNWELL, JOHN
APPLICANT: MENDIS, KAMINI
TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF
TITLE OF INVENTION: PLASMODIUM MSP-1
FILE REFERENCE: 0660-0135-OXIP
CURRENT APPLICATION NUMBER: US/09/134,333
CURRENT FILING DATE: 1999-04-18
EARLIER APPLICATION NUMBER: PCT/FR97/00290
EARLIER FILING DATE: 1997-02-14
EARLIER APPLICATION NUMBER: FR96/01822
EARLIER FILING DATE: 1996-02-14
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 11
LENGTH: 379
TYPE: PRT
ORGANISM: Plasmodium cynomolgi
FEATURE:
OTHER INFORMATION: Amino Acids 1-139- REGION I
FEATURE:
OTHER INFORMATION: Amino Acids 140-177-REGION II
FEATURE:
OTHER INFORMATION: Amino Acids 178-282-REGION III
FEATURE:
OTHER INFORMATION: Amino Acids 283-379-REGION IV
US-09-134-333-11

Alignment Scores:
Pred. No.:           0.302           Length:           379
Score:              105.50          Matches:            80
Percent Similarity:  40.62%         Conservative:       63
Best Local Similarity: 22.73%        Mismatches:       127
Query Match:         2.95%          Indels:            82
DB:                  10             Gaps:              14

US-09-513-151-3 (1-2041) x US-09-134-333-11 (1-379)
Oy      158  GAGATCGTAGCGGTACCTCCATGCAGAGTGTCATGAAGCCTAGACATCATCACCAACAAG 217
          |||||  |||  ::|||  ||||||:::
Db      14   Guillellevalproclngllyleasnglutyf-----AspValaITYrllelys 30
Oy      218  GTTCTGCCCAAGAGCAGAATCTCCGGCACACATGATCAGCTTGATGGATCCCTCTT 277
          |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db      31  ProleuaAlaglymetyrlsthrilleylsLysProleuclunshlslalaleu 50
Oy      278  GTGACCAAAT--TACACAGTGTGAGACTTCAGAAATAGAACCACTGCTCGATTGAAGAT 334
          |||||  ::::::|||  |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db      51  AsnThrAsnllleleaspmetleuaspsarArgleuLyslsyrsArgsanTyrrPheleuAsp 70
Oy      335  ATATTTCGCCGAGACAAAATTCCTATTGTTGGGAGGAACAATTAATTCATTGAATCT 394
          ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db      71  ValLeuAsnSerAspleuasnProtyrserIlleProHisserIlglyutyrllellelys 90
Oy      395  CTGCTGTGAAAGTCTTGTCAATACCAAGCCCCAGAGAGTGGCACTGAGAAAGTATT 454
          ::::::|||  |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db      91  AspProTyrglyleuasnspLeugluLysLysleuendulserTyrlvstyrylle 110
Oy      455  GACCGAAAAAGTGAG-----CTTGAAGAGGAGATGCTTGACTTAC----- 499
          ||:::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db      111  GlyAlaserValaspysaspmetvalThrrlaAsnAspolyleuallaryTyrglnLys 130
          :::::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Oy      500  -----AACGCCTAAGCCAGGTGAC-----CCAGAAATGCGT 532

```

```

Db      131 MetGlyaspLeuTYrLysIleHisLeuAspGluValAlaSnalAcylSileYsgIuValGlu 150
QY      533 GCCAAGCTGCATCCACATGAC-----                    553
          ||| :||| | |||||
Db      151 AlaSnIleasnLYshISasPGLUGlUlleYslSLieGlyserGIuaIserySaLa 170
QY      554 -----AAACGCCAAAGTGCCCGACGACTTGCAGATTTT 568
          ||| :||| | |||||
Db      171 AsnaSPlysASngInLeuAsnAlaLySLySGlUGlueNGlnLYstYrLeu- ProPheLe 190
QY      587 G-----AAGAACAAGCATCTTCATGTGAAT-----TTTCACATCGCACAT 631
          | :||| | |||||
Db      190 uSerIErliegInLYsgIUtyrSErThrLeuValAsnLYsValHiseRYthrASPth 210
QY      632 ACGGAAGAAGGTGGTGTGCCCCCTTGAGAGTCCCTCGAATCTTCAACCCTTCATCCTT 691
          ||||| :||| | |||||
Db      210 rLeuLYslYslIErlleIASnaNCySGlnIle- GluLYslYsgIUtnrGlutHlleVal 229
QY      692 TGCGTTCATCGCTGCAGACCGACGAGTCTTAATAGCGCTTGAT-----AAG 736
          ||| :||| | |||||
Db      230 ASnLYslSeuGluASpTYrSErLYseRysMetAspGluGluLeuAspValTYrLYsgInSERlys 249
QY      737 AGGGTGGATGCATCCTTGCTGCTGGCGGCTCTTGAGGAGAACATAAGAT-----TTT 787
          :||| :||| | |||||
Db      250 LYsgIuAspAspValLYsgISErSErCYleuENGluLYslLeuMetAsnSERlysLeuIle 265
QY      788 CACACAGCCCTTAATACAGAAATAAGTTTTCGAA-----AATAGCAGACTAT 835
          :||| :||| | |||||
Db      270 ASngInGUglUserLYslYsALeUsErCluLeuLeuAsnValGInTHgIn-MetLe 289
QY      836 CAACATGCTATCTCCAATCAATTAGGCTTCAAGCAATTTACGAGTACCTGATCTACAG 895
          ||||| :||| | |||||
Db      289 uASnMEtSErSErCluHISArCySlleAspThrASN-----ValPr 303
QY      896 GGAAATATCA-----CACTGGAGACTAGTAACAG 922
          ||||| :||| | |||||
Db      303 GdluSnMaIlaaLYcSYrFArGYrLEuAspGlyTHgIGluGluTrpArgCYslEuleTY 323
QY      926 CTTCTTAAGAAAGACCTCGTCCCATGTGCTCTGTATGCGCTTACAGATTAATCTGAT 985
          ||| ||||| :||| | |||||
Db      323 rPhE-LySGlu-----AspAlaGlyLYscYsVALProAlaProAsnMeThrc 339
QY      986 GTCTCGAAGTGGAGGAGCTGCTTCT 1011
          || :||| | |||||
Db      339 ySLySAspLYsASnGlyCYsAla 347

RESULT 12
US-09-792-630-41
; Sequence 41, Application US/09792630
; Patent No. US20020168640A1
; GENERAL INFORMATION:
; APPLICANT: LI, Min
; APPLICANT: Dahljat, Bassil I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/792.630
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 41
; LENGTH: 671
; TYPE: PRT
; ORGANISM: B19 virus
US-09-792-630-41

Alignment Scores:
Pred. NO.: 0.668 Length: 671
Score: 103.00 Matches: 116
Percent Similarity: 33.89% Conservative: 67
Best Local Similarity: 21.48% Mismatches: 170
Query Match: 2.82% Indels: 187
DB: 9 Gaps: 30

```


OY	1799	TCACTGACATTCAGCATCCCAATATGTGGCGTCCACAATAATTACTTGCTATCCTACAAAAGC	1746
Db	213	SerIleLySerheciInthrMetValAsnTrpLeuCysGIuAsnArgValPheThcGIuSp	232
OY	1739	ATCTGGAAATTAGATTAATTTTAAACCAG-----AGTCAGGAGCATPAAACTTCT	1692
Db	233	LysTrpLysLeuValAsnPheAsnGIuNryrThrLeuLeuSerSetSerHisSerGIySer	252
OY	1691	TTTTAAAGGATGCAGCTCAATCTCG-----GTAATTCAAC	1659
Db	253	PheGIu----IleGIuSerAlaLeuLysLeuAlaIleTyGLysAlaThrAsnLeuValPro	271
OY	1598	TTAAGGTTCAAAGAAAAAATGCTTTCCG-----AACATAC-----	1566
Db	288	-----LysAspAsnLysIleValLysLeuLeuLysGIuAsnTrpProLeu	304
OY	1559	ATCATTTCCAGACACATCACGCCACACAGAAGAGCTGCACAAGAGCTGCTTTCTATTATAG	1500
Db	305	LeuValGIuGIuHisVal-----	310
OY	1499	AGAACGTGAGACTTTAAACACATCAAAAAGAAATGTGGAGCTTTTCGTCTATGCAG	1440
Db	311	-----LeuLysTrpIleAspLys-----	316
OY	1439	AGATTCCGCATAGCATCTCTTGCCACACATGGGAGGACAACATATAC-----	1392
Db	317	-----LysCysGIuLysLysAsnThrLeuTrpPheTyR	327
OY	1391	---CTTCCTCTCGAATCGATCGATCCCACACACCTTTCCAAAGGCCATGTCTCTT	1335
Db	328	GlyProProSer---Thrgly-----LysThrAsnLeuAlaMetAlaIle	341
OY	1334	AAMCGCGCATTTACAGCTCTGTATGATTCATTCGGCCGGGAGATCCCTCCCTTAGGTC-	1277
Db	342	AlaLysSerValProValTyGIu-MetValAsnTrpAsnAsnGIuAsnPheProPheAs	361
OY	1276	-----TTTGTTATAGCTCGGGAAACATCTGCAGCTTCTATAGCTGTGA	1233
Db	361	nAspValAlaGIuLysSerLeuValAlrAspGIuGIuIleIle-----	376
OY	1232	CAGCATCTGAGTCCAATCTTCTTTCTTCAAGTTGGTCAAGTGGCATTTGGATTMTTA	1173
Db	377	-----LysSerThrIleValGIuAlaIleAlaLysAlaIleLeuGIuGIyGI	391
OY	1172	T-----GTGGCGGCCCATCGCGATGCCCATGATGATTCAT	1134
Db	391	nProThrArgValAspGIuLysMetArgGIuSerValAlaValProGIu-----	407
OY	1133	CACAGAGTCCACAGATGATTAATCTCTCT--GTTCCAGATTCATTTGATGCATCT	1077
Db	408	-----ValProValIleThrSerAsnGIuLysPheValIleSerGIyAs	425
OY	1076	TTATTTGAGTGGCTGTAAGCTGTGGCCCTGTGATGAATTTGCACAGATTTCAAGAGCAG	1017
Db	425	nThrThrTrpThrValHisAlaLysAlaLeuLysGIuArgMet-----ValLysLeuAs	443
OY	1016	GTTTCAACAAGACTCTCCCA-----CTTCCAGACATCAGATACCTCTTAAGC	969
Db	443	nPheThrValArGIuCySerProAspMetGIuLeuLeuTrgIu-AlaAspValGIuIngIT	463
OY	968	CATAGACAGGGGAGCAATGGACAGGCTCTTTCTTTTGAAGCTGTACTAGTCTCCA	909
Db	463	rPheUThrTrpCysAsnAlaGIuInsErTrpAspHisTrGIuAsnTrpAlaIleAsnTyT	483
OY	908	GTCGTGACTTTTCCCTCAAGTGATCAGGTAAGTCTGTGAATTCCTTTGAAGCCAATTCATTGGA	849

```

Db      483 hrhpharpherrgclyle-----asnlaasrallaleuinhspdrleuclnt 500
Qy      848 AGATASCATGTTGATAGTCTCTGCTATTTTCCGAACATTTCTGTGATTAGCGCTGT 789
Db      500 hrthrPro-----llevalthrAspThr----- 507
Qy      788 GAAATCGCTAGTTCCTCCCAAGAGCCSAGAGACCATGTATCCACCCCTTATGCCA 729
Db      508 -----SerIleSerSerSer-----GlyIlyGluSerSerGluGluLeuSerG 522
Qy      728 AGCGCTCA-----TCTAGAACCTCGCTGTCAGCATGCAAGCCAAAGCATGC 684
Db      522 luserSerPhephasnleuIethrProGlyAlatrpaSnThr----- 536
Qy      683 AAGGATTAGAACCTTCAGAGACCTCCAAAGGGAGCACCACCTTCTTCGCTATGTTGAC 624
Db      537 -----GluThrProArgSerSerThrPro-IleProGlyThrSer 549
Qy      623 GATGAGAGAAATTCACATAGAGATTCCTGTTTTCATAAAATCTTGACACTCTGCG-- 567
Db      550 SerGlyGluSerPheValGlySerSerValSerSerGluValAlaIalaSerTrpGlu 569
Qy      566 -----CCACTTTCG-----GTTTCGTCATGTGGATGCA 540
Db      570 GluAlaIethrTrhTrProIleAlaAlaAspGlnPheArgGluLeuLeuValGlyValAspTr 589
Qy      539 GCTTGCGAGCCATTTCTGGGTCSCACCTGCGCTTAGCGCTTGTGTAAGTACAAGACCATCT 480
Db      590 ValTrpAsp-----GlyValArgGlyLeuProValAlaCysCysValGln----- 603
Qy      479 CCTTTTCAAGCTCCACTTTTGGGTCATCACTTTCTAGTGGCCA-----TCTCCT 429
Db      604 -----HisIleasnAsnSerGlyGlyGlyLeuGlyLeuGlyCysProHisCysIleasVal 621
Qy      428 GGGCGCTGATTGTGACAAGACTTTCCAGAGACAGATTCATATGTAATTAATGCG 375
Db      622 GlyAlatrTrpIlyr-----asnGlyTrp 628

RESULT 13
US-10-080-376-41
Sequence 41, Application US/10080376
Patent NO. US20020172968A1
GENERAL INFORMATION:
APPLICANT: Li, Min
TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
FILE REFERENCE: A-70295-2/RT/RMS/RMK
CURRENT APPLICATION NUMBER: US/10/080, 376
CURRENT FILING DATE: 2000-02-19
PRIOR APPLICATION NUMBER: US 09/792, 630
PRIOR FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 87
SOFTWARE: PatentIn version 3.1
SEQ ID NO 41
LENGTH: 671
TYPE: PRT
ORGANISM: B19 virus
US-10-080-376-41

```

Alignment Scores:	
Pred. No.:	0.688
Score:	103.00
Percent Similarity:	33.89%
Best Local Similarity:	21.48%
Query Match:	2.82%
DB:	9
US-09-513-151-3 (1-2041) x US-10-080-376-41 (1-671)	
Length:	671
Matches:	116
Conservative:	67
Mismatches:	170
Indels:	187
Gaps:	30

0y 1799 TCACCAATATTAGAACTCCCAATATGTGGCTCAACAATCTTCAGTCGATCTCAAAAGC 174
||||:|||| :||||| ||||| ||| :||
Db 213 SerIeLySPheGlnThrMeValAsnTrpLeuYscIuAsnArGValPheThrInGluAsp 232

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 19:18:40 ; Search time 37.5 Seconds

(without alignments)
8230.823 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 661
Sequence: 1 CTCGCATACATGCGTCGCG.....TTTACAGAGAAAAA 2041

Scoring table:

OLIGO	Xgapop 60.0	Xgapext 60.0
	Ygapop 60.0	Ygapext 60.0
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 288829 seqs, 75613885 residues

Word size: 1

Total number of hits satisfying chosen parameters: 541640

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODEL=frame+ol2p.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09513151/runat_15042003_141249_27326/app.query.fasta_1.2183
-DB=published_applications_AA -OFMT=fastan -SUFFIX=ol1n2p.rapb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=dots -START=1 -END=1 -MATRIX=ol1n2p
-TRANS=human40.cdi -LIST=45 -DICALIGN=200 -THR_SCORE=quality -THR_MIN=1
-MAXLEN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09513151 -ECGN_1_1_13 -erunat_15042003_141249_27326
-NCPU=6 -ICPU=3 -NO_XLPTX -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FCGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database: Published Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PC1_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PC105_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	123	18.6	222	US-09-764-853-586	Sequence 586, App
2	123	18.6	222	US-09-764-853-784	Sequence 784, App
3	48	7.3	57	US-09-864-761-42725	Sequence 42725, A
4	8	1.2	115	US-09-867-550-1444	Sequence 1444, Ap

	5	8	1.2	139	10	US-09-867-550-482	Sequence 482, App
	6	8	1.2	378	9	US-09-970-989-17	Sequence 17, Appl
	7	8	1.2	378	9	US-09-992-598-156	Sequence 156, App
	8	8	1.2	378	9	US-09-989-293A-156	Sequence 156, App
	9	8	1.2	378	9	US-09-988-725-156	Sequence 156, App
	10	8	1.2	378	9	US-09-990-444-156	Sequence 156, App
	11	8	1.2	378	9	US-09-989-730-156	Sequence 156, App
	12	8	1.2	378	9	US-09-990-436-156	Sequence 156, App
	13	8	1.2	378	9	US-09-991-181-156	Sequence 156, App
	14	8	1.2	378	9	US-09-993-667-156	Sequence 156, App
	15	8	1.2	378	9	US-09-988-724-156	Sequence 156, App
	16	8	1.2	378	9	US-10-028-072-420	Sequence 420, App
	17	8	1.2	378	9	US-09-997-653-156	Sequence 156, App
	18	8	1.2	378	9	US-09-993-667-156	Sequence 156, App
	19	8	1.2	378	9	US-10-121-049-420	Sequence 420, App
	20	8	1.2	378	9	US-10-123-904-420	Sequence 420, App
	21	8	1.2	378	9	US-10-140-470-420	Sequence 420, App
	22	8	1.2	378	9	US-09-990-438-156	Sequence 156, App
	23	8	1.2	378	9	US-09-990-562-156	Sequence 156, App
	24	8	1.2	378	9	US-09-997-428-156	Sequence 156, App
	25	8	1.2	378	9	US-09-997-666-156	Sequence 156, App
	26	8	1.2	378	9	US-10-175-746-420	Sequence 420, App
	27	8	1.2	378	9	US-10-176-918-420	Sequence 420, App
	28	8	1.2	378	9	US-10-176-921-420	Sequence 420, App
	29	8	1.2	378	9	US-09-990-711-156	Sequence 156, App
	30	8	1.2	378	9	US-10-137-865-420	Sequence 420, App
	31	8	1.2	378	9	US-10-140-474-420	Sequence 420, App
	32	8	1.2	378	9	US-10-142-431-420	Sequence 420, App
	33	8	1.2	378	9	US-10-143-114-420	Sequence 420, App
	34	8	1.2	378	9	US-09-989-726-156	Sequence 156, App
	35	8	1.2	378	9	US-10-140-002-420	Sequence 420, App
	36	8	1.2	378	9	US-09-990-437-156	Sequence 156, App
	37	8	1.2	378	9	US-09-998-156-156	Sequence 156, App
	38	8	1.2	378	9	US-10-142-419-420	Sequence 420, App
	39	8	1.2	378	9	US-09-991-157-156	Sequence 156, App
	40	8	1.2	378	9	US-09-991-172-156	Sequence 156, App
	41	8	1.2	378	9	US-09-997-513-156	Sequence 156, App
	42	8	1.2	378	9	US-09-997-574-156	Sequence 156, App
	43	8	1.2	378	9	US-10-123-252-420	Sequence 420, App
	44	8	1.2	378	9	US-10-142-423-420	Sequence 420, App
	45	8	1.2	378	9	US-09-990-443-156	Sequence 156, App

ALIGNMENTS

RESULT 1
US-09-764-853-586
Sequence 586, Application US/09764853
Patent No. US20020090672A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PJ206
CURRENT APPLICATION NUMBER: US/09/764,853
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 939
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 586
LENGTH: 222
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (124)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (145)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-586
Alignment Scores:
Pred. No.: 2.12e-112 Length: 222

Score: 123.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.61% Indels: 0
DB: 10 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-764-853-784 (1-222)

QY 2 TGGCATTAGATGGCGTCCGTGGCGGCTGCACAGACACTTCTGTGGCAGTGGGCTCAGG 61
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
QY 62 GGGCTGCACGAGACCCCTACCTCTTGTAGTATCTGGGGCCACGGGACCGGCAATTC 121
Db 21 GlyLeuGlnArgThrLeuProLeuValAlaIleLeuGlyAlaThrGlyThrGlyLysSer 40
QY 122 AGCGTGGCTGTGAGCTAGAGCCAGCGGCTCGGCGGTGAGATCGTCACGCTGATCCATG 181
Db 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGluIleValSerAlaAspSerMet 60
QY 182 CAGGCTATAGAGGCTTACATCATCATCACCACAGGTTTCTGCCACAGAGAGATC 241
Db 61 GlnValTyrluGluGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80
QY 242 TGGCGGACCAACATGATCAGCTTGTGATCTCTGTGACCAATTACAGAGGTGAGC 301
Db 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrlThrValAlaAsp 100
QY 302 TTCGAAATAGAGCACTGCTGTGATTGAAGATATATTGCCGAGCAAAATTCTTATT 361
Db 101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle 120
QY 362 GTTGTGGGA 370
Db 121 ValValGly 123

RESULT 2
US-09-764-853-784
; Sequence 784, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P2206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 784
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-784

Alignment Scores: 2.12e-112 Length: 222
Score: 123.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.61% Indels: 0
DB: 10 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-764-853-784 (1-222)

QY 2 TGGCATTAGATGGCGTCCGTGGCGGCTGCACAGACACTTCTGTGGCAGTGGGCTCAGG 61
2 TGCATTAAGATGGCGTCCGTGGCGGCTGCACAGACACTTCTGTGGCAGTGGGCTCAGG 61
1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
62 GGGCTGCACGAGACCCCTACCTCTTGTAGTATCTGGGGCCACGGGACCGGCAATTC 121
21 GlyLeuGlnArgThrLeuProLeuValAlaIleLeuGlyAlaThrGlyThrGlyLysSer 40
122 AGCGTGGCTGTGAGCTAGAGCCAGCGGCTCGGCGGTGAGATCGTCACGCTGATCCATG 181
41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGluIleValSerAlaAspSerMet 60
182 CAGGCTATAGAGGCTTACATCATCATCACCACAGGTTTCTGCCACAGAGAGATC 241
61 GlnValTyrluGluGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80
242 TGGCGGACCAACATGATCAGCTTGTGATCTCTGTGACCAATTACAGAGGTGAGC 301
81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrlThrValAlaAsp 100
302 TTCGAAATAGAGCACTGCTGTGATTGAAGATATATTGCCGAGCAAAATTCTTATT 361
101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle 120
362 GTTGTGGGA 370
121 ValValGly 123

Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
QY 62 GGGCTGCACGAGACCCCTACCTCTTGTAGTATCTGGGGCCACGGGACCGGCAATTC 121
Db 21 GlyLeuGlnArgThrLeuProLeuValAlaIleLeuGlyAlaThrGlyThrGlyLysSer 40
QY 122 AGCGTGGCTGTGAGCTAGAGCCAGCGGCTCGGCGGTGAGATCGTCACGCTGATCCATG 181
Db 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGluIleValSerAlaAspSerMet 60
QY 182 CAGGCTATAGAGGCTTACATCATCATCACCACAGGTTTCTGCCACAGAGAGATC 241
Db 61 GlnValTyrluGluGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80
QY 242 TGGCGGACCAACATGATCAGCTTGTGATCTCTGTGACCAATTACAGAGGTGAGC 301
Db 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrlThrValAlaAsp 100
QY 302 TTCGAAATAGAGCACTGCTGTGATTGAAGATATATTGCCGAGCAAAATTCTTATT 361
Db 101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle 120
QY 362 GTTGTGGGA 370
Db 121 ValValGly 123

RESULT 3
US-09-864-761-42725
; Sequence 42725, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 42725
LENGTH: 57
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL033527.25
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: EST_HUMAN HIT: BE242161.1, EVALUATE 3.00e-21
OTHER INFORMATION: SWISSPROT HIT: Q9KAC3, EVALUATE 2.00e-04
US-09-864-761-42725

Alignment Scores:
Pred. No.: 2,51e-38 Length: 57
Score: 48.00 Matches: 48
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.26% Indels: 0
DB: 10 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-864-761-42725 (1-57)

OY 185 GTCATGAGCGCTAGACATATCCACACAGGTTTGGCCACAGACAGAGATCTGC 244
|||||
Db 1 VAlTyrgIuglYleuAsPlleIleThrAsnYsValSerAlaGInGInGInArG11eCys 20

OY 245 CGGACACGATGATCAGCTTGTGAGACCTCTTGACACATTCACAGTGGAGCTTC 304
|||||
Db 21 ArGhshShMcIlleSePheValAsPProLeuValThrAsnYrYThValValAsPhe 40

OY 305 AGAATAGAGCACTGCTGATT 328
|||||
Db 41 ArGAsnArGAlaThrAlaLeuIle 48

RESULT 4
US-09-867-550-1444

Sequence 1444, Application US/09867550
Patent No. US20020082206A1

GENERAL INFORMATION:

APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Fuad,
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie
APPLICANT: Topper, James
TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
TITLE OF INVENTION: Thereby

FILE REFERENCE: 21402-013 (Cura-313)

CURRENT APPLICATION NUMBER: US/09/867,550

CURRENT FILING DATE: 2001-09-20

PRIOR APPLICATION NUMBER: USSN 60/208,427

PRIOR FILING DATE: 2000-05-30

NUMBER OF SEQ ID NOS: 2125

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1444

LENGTH: 115

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)

OTHER INFORMATION: Wherein Xaa may be any one of Arg or Gly or Trp
US-09-867-550-1444

Alignment Scores:
Pred. No.: 70.2 Length: 115
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.22% Indels: 0
DB: 10 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-867-550-1444 (1-115)

OY 811 ATTCTTGATTAAGCCTGCTG 788
|||||
Db 2 lIleuLeuIleIleAlaSerVal 9

RESULT 5
US-09-867-550-482

Sequence 482, Application US/09867550
Patent No. US20020082206A1

GENERAL INFORMATION:

APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Fuad,
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie
APPLICANT: Topper, James
TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cell
TITLE OF INVENTION: Thereby

FILE REFERENCE: 21402-013 (Cura-313)

CURRENT APPLICATION NUMBER: US/09/867,550

CURRENT FILING DATE: 2001-09-20

PRIOR APPLICATION NUMBER: USSN 60/208,427

PRIOR FILING DATE: 2000-05-30

NUMBER OF SEQ ID NOS: 2125

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 482

LENGTH: 139

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

Alignment Scores:
Pred. No.: 69 Length: 139
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 10 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-867-550-482 (1-139)

OY 113 GGCAATCCAGCGCTGCTGCAG 136
|||||
Db 14 GlYlYsSerThrIleuAlaLeuGln 21

RESULT 6
US-09-970-989-17

Sequence 17, Application US/09970989
Patent No. US20020156262A1

GENERAL INFORMATION:

APPLICANT: LEUNG, David W.
APPLICANT: ADONREI, Daniel
APPLICANT: HOLLENBACK, David
TITLE OF INVENTION: MAMMALIAN LYOPHOSPHATIDIC ACID ACYLTRANSFERASE
TITLE OF INVENTION: 077319/0151
FILE REFERENCE: 077319/0151
CURRENT APPLICATION NUMBER: US/09/970,989
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 09/215,252
PRIOR FILING DATE: 1998-12-18
PRIOR APPLICATION NUMBER: US 08/618,651
PRIOR FILING DATE: 1996-03-19

NUMBER OF SEQ ID NOS: 42

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 17

```

; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-989-17

Alignment Scores:
Pred. No.: 62.8 Length: 378
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 9 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-970-989-17 (1-378)
OY 669 AGTTCCTGAACCTGCATCTT 692
Db 334 SerSerLeuThrLeuAlaSerPhe 341

RESULT 7
US-09-992-598-156
; Sequence 156, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Bolstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerbtsen, Hanspeter
; APPLICANT: Goddard, Audrey E.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC20
; CURRENT APPLICATION NUMBER: US/09/992,598
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
```

```

; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
```


;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07

;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 62.8
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.21%
DB: 9
Gaps: 0

US-09-513-151-3 (1-2041) x US-09-992-598-156 (1-378)
OY 669 AGTCTCTAACCTTCATCCTT 692
|||||
Db 334 SerSerLeuThrLeuAlaSerPhe 341

RESULT 8
US-09-989-293A-156
Sequence 156, Application US/09989293A
Patent No. US20020177164A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavini, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Collin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhanq, Zemin
TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C6
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: US/09/989, 293A
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02

1 PRIOR APPLICATION NUMBER: 60/087609
2 PRIOR FILING DATE: 1998-06-02
3 PRIOR APPLICATION NUMBER: 60/087759
4 PRIOR FILING DATE: 1998-06-02
5 PRIOR APPLICATION NUMBER: 60/087827
6 PRIOR FILING DATE: 1998-06-03
7 PRIOR APPLICATION NUMBER: 60/088021
8 PRIOR FILING DATE: 1998-06-04
9 PRIOR APPLICATION NUMBER: 60/088025
10 PRIOR FILING DATE: 1998-06-04
11 PRIOR APPLICATION NUMBER: 60/088026
12 PRIOR FILING DATE: 1998-06-04
13 PRIOR APPLICATION NUMBER: 60/088028
14 PRIOR FILING DATE: 1998-06-04
15 PRIOR APPLICATION NUMBER: 60/088029
16 PRIOR FILING DATE: 1998-06-04
17 PRIOR APPLICATION NUMBER: 60/088030
18 PRIOR FILING DATE: 1998-06-04
19 PRIOR APPLICATION NUMBER: 60/088033
20 PRIOR FILING DATE: 1998-06-04
21 PRIOR APPLICATION NUMBER: 60/088326
22 PRIOR FILING DATE: 1998-06-04
23 PRIOR APPLICATION NUMBER: 60/088167
24 PRIOR FILING DATE: 1998-06-05
25 PRIOR APPLICATION NUMBER: 60/088202
26 PRIOR FILING DATE: 1998-06-05
27 PRIOR APPLICATION NUMBER: 60/088212
28 PRIOR FILING DATE: 1998-06-05
29 PRIOR APPLICATION NUMBER: 60/088217
30 PRIOR FILING DATE: 1998-06-05
31 PRIOR APPLICATION NUMBER: 60/088655
32 PRIOR FILING DATE: 1998-06-09
33 PRIOR APPLICATION NUMBER: 60/088734
34 PRIOR FILING DATE: 1998-06-10
35 PRIOR APPLICATION NUMBER: 60/088738
36 PRIOR FILING DATE: 1998-06-10
37 PRIOR APPLICATION NUMBER: 60/088742
38 PRIOR FILING DATE: 1998-06-10
39 PRIOR APPLICATION NUMBER: 60/088810
40 PRIOR FILING DATE: 1998-06-10
41 PRIOR APPLICATION NUMBER: 60/088824
42 PRIOR FILING DATE: 1998-06-10
43 PRIOR APPLICATION NUMBER: 60/088826
44 PRIOR FILING DATE: 1998-06-10
45 PRIOR APPLICATION NUMBER: 60/088858
46 PRIOR FILING DATE: 1998-06-11
47 PRIOR APPLICATION NUMBER: 60/088861
48 PRIOR FILING DATE: 1998-06-11
49 PRIOR APPLICATION NUMBER: 60/088876
50 PRIOR FILING DATE: 1998-06-11
51 PRIOR APPLICATION NUMBER: 60/089105
52 PRIOR FILING DATE: 1998-06-12
53 PRIOR APPLICATION NUMBER: 60/089440
54 PRIOR FILING DATE: 1998-06-16
55 PRIOR APPLICATION NUMBER: 60/089512
56 PRIOR FILING DATE: 1998-06-16
57 PRIOR APPLICATION NUMBER: 60/089514
58 PRIOR FILING DATE: 1998-06-16
59 PRIOR APPLICATION NUMBER: 60/089532
60 PRIOR FILING DATE: 1998-06-17
61 PRIOR APPLICATION NUMBER: 60/089538
62 PRIOR FILING DATE: 1998-06-17
63 PRIOR APPLICATION NUMBER: 60/089598
64 PRIOR FILING DATE: 1998-06-17
65 PRIOR APPLICATION NUMBER: 60/089599
66 PRIOR FILING DATE: 1998-06-17
67 PRIOR APPLICATION NUMBER: 60/089600
68 PRIOR FILING DATE: 1998-06-17
69 PRIOR APPLICATION NUMBER: 60/089653
70 PRIOR FILING DATE: 1998-06-17
71 PRIOR APPLICATION NUMBER: 60/089801
72 PRIOR FILING DATE: 1998-06-18
73 PRIOR APPLICATION NUMBER: 60/089907
74 PRIOR FILING DATE: 1998-06-18
75 PRIOR APPLICATION NUMBER: 60/089908
76 PRIOR FILING DATE: 1998-06-18
77 PRIOR APPLICATION NUMBER: 60/089947
78 PRIOR FILING DATE: 1998-06-19
79 PRIOR APPLICATION NUMBER: 60/089948
80 PRIOR FILING DATE: 1998-06-19
81 PRIOR APPLICATION NUMBER: 60/089952
82 PRIOR FILING DATE: 1998-06-19
83 PRIOR APPLICATION NUMBER: 60/090246
84 PRIOR FILING DATE: 1998-06-22
85 PRIOR APPLICATION NUMBER: 60/090252
86 PRIOR FILING DATE: 1998-06-22
87 PRIOR APPLICATION NUMBER: 60/090254
88 PRIOR FILING DATE: 1998-06-22
89 PRIOR APPLICATION NUMBER: 60/090349
90 PRIOR FILING DATE: 1998-06-23
91 PRIOR APPLICATION NUMBER: 60/090355
92 PRIOR FILING DATE: 1998-06-23
93 PRIOR APPLICATION NUMBER: 60/090429
94 PRIOR FILING DATE: 1998-06-24
95 PRIOR APPLICATION NUMBER: 60/090431
96 PRIOR FILING DATE: 1998-06-24
97 PRIOR APPLICATION NUMBER: 60/090435
98 PRIOR FILING DATE: 1998-06-24
99 PRIOR APPLICATION NUMBER: 60/090444
100 PRIOR FILING DATE: 1998-06-24
101 PRIOR APPLICATION NUMBER: 60/090445
102 PRIOR FILING DATE: 1998-06-24
103 PRIOR APPLICATION NUMBER: 60/090472
104 PRIOR FILING DATE: 1998-06-24
105 PRIOR APPLICATION NUMBER: 60/090535
106 PRIOR FILING DATE: 1998-06-24
107 PRIOR APPLICATION NUMBER: 60/090540
108 PRIOR FILING DATE: 1998-06-24
109 PRIOR APPLICATION NUMBER: 60/090542
110 PRIOR FILING DATE: 1998-06-24
111 PRIOR APPLICATION NUMBER: 60/090557
112 PRIOR FILING DATE: 1998-06-24
113 PRIOR APPLICATION NUMBER: 60/090676
114 PRIOR FILING DATE: 1998-06-25
115 PRIOR APPLICATION NUMBER: 60/090678
116 PRIOR FILING DATE: 1998-06-25
117 PRIOR APPLICATION NUMBER: 60/090690
118 PRIOR FILING DATE: 1998-06-25
119 PRIOR APPLICATION NUMBER: 60/090694
120 PRIOR FILING DATE: 1998-06-25
121 PRIOR APPLICATION NUMBER: 60/090695
122 PRIOR FILING DATE: 1998-06-25
123 PRIOR APPLICATION NUMBER: 60/090696
124 PRIOR FILING DATE: 1998-06-25
125 PRIOR APPLICATION NUMBER: 60/090862
126 PRIOR FILING DATE: 1998-06-26
127 PRIOR APPLICATION NUMBER: 60/090863
128 PRIOR FILING DATE: 1998-06-26
129 PRIOR APPLICATION NUMBER: 60/091360
130 PRIOR FILING DATE: 1998-07-01
131 PRIOR APPLICATION NUMBER: 60/091478
132 PRIOR FILING DATE: 1998-07-02
133 PRIOR APPLICATION NUMBER: 60/091544
134 PRIOR FILING DATE: 1998-07-01
135 PRIOR APPLICATION NUMBER: 60/091519
136 PRIOR FILING DATE: 1998-07-02
137 PRIOR APPLICATION NUMBER: 60/091626
138 PRIOR FILING DATE: 1998-07-02
139 PRIOR APPLICATION NUMBER: 60/091633
140 PRIOR FILING DATE: 1998-07-02
141 PRIOR APPLICATION NUMBER: 60/091978
142 PRIOR FILING DATE: 1998-07-07
143 PRIOR APPLICATION NUMBER: 60/091982
144 PRIOR FILING DATE: 1998-07-07
145 PRIOR APPLICATION NUMBER: 60/092182
146 PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. NO.:	62.8	Length:	378
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	9	Gaps:	0

US-09-513-151-3 (1-2041) x US-09-989-293A-156 (1-378)

OY 669 AGTCTCTAACCTTGATCCTT 692
|||||

Db 334 SerSerLeuThrLeuAlaSerPhe 341

RESULT 9

US-09-989-735-156

Sequence 156, Application US/09989735

Publication No. US20020193299A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Bolstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, J. Christopher

APPLICANT: Guiney, Austin L.

APPLICANT: Kljavin, Ivar J.

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

TITLE OF INVENTION: Acids Encoding the Same

FILE REFERENCE: P2730P1C61

CURRENT APPLICATION NUMBER: US/09/989,735

CURRENT FILING DATE: 2001-11-19

PRIOR APPLICATION NUMBER: 60/049787

PRIOR FILING DATE: 1997-06-16

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/065186

PRIOR FILING DATE: 1997-11-12

PRIOR APPLICATION NUMBER: 60/065311

PRIOR FILING DATE: 1997-11-13

PRIOR APPLICATION NUMBER: 60/066770

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/075945

PRIOR FILING DATE: 1998-02-25

PRIOR APPLICATION NUMBER: 60/078910

PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/083322

PRIOR FILING DATE: 1998-04-28

PRIOR APPLICATION NUMBER: 60/084600

PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/087106

PRIOR FILING DATE: 1998-05-28

PRIOR APPLICATION NUMBER: 60/087607

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087609

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087759

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087827

PRIOR FILING DATE: 1998-06-03

PRIOR APPLICATION NUMBER: 60/088021

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088025

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088026

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088028

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088029

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088030

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088033

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088326

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088167

PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/088202

PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/088212

PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/088217

PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/088655

PRIOR FILING DATE: 1998-06-09

PRIOR APPLICATION NUMBER: 60/088734

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088738

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088742

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088810

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088824

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088826

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088858

PRIOR FILING DATE: 1998-06-11

PRIOR APPLICATION NUMBER: 60/088861

PRIOR FILING DATE: 1998-06-11

PRIOR APPLICATION NUMBER: 60/088876

PRIOR FILING DATE: 1998-06-11

PRIOR APPLICATION NUMBER: 60/089105

PRIOR FILING DATE: 1998-06-12

PRIOR APPLICATION NUMBER: 60/089440

PRIOR FILING DATE: 1998-06-16

PRIOR APPLICATION NUMBER: 60/089512

PRIOR FILING DATE: 1998-06-16

PRIOR APPLICATION NUMBER: 60/089514

PRIOR FILING DATE: 1998-06-16

PRIOR APPLICATION NUMBER: 60/089532

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/089538

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/089598

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/089599

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/089600

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/089653

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/089801

PRIOR FILING DATE: 1998-06-18

PRIOR APPLICATION NUMBER: 60/089907

PRIOR FILING DATE: 1998-06-18

PRIOR APPLICATION NUMBER: 60/089908

PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.: 62.8 Length: 378
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 9 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-989-735-156 (1-378)

OY 669 AGTTCCTTAACCCCTGCATCCTTT 692
|||||
Db 334 SerSerleuthreunAlaserphe 341

RESULT 10
US-09-990-444-156
Sequence 156, Application US/09990444
Publication No. US20020193300A1

GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumms, Daniel
APPLICANT: Watanabe, Collin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC19
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02


```
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.21%                Indels: 0
DB: 9                              Gaps: 0
US-09-513-151-3 (1-2041) x US-09-990-444-156 (1-378)

QY 669 AGTTCCTACCTGCATCCTT 692
DB 334 SerSetleuThrLeuAlaSerPhe 341

RESULT 11
US-09-989-730-156
: Sequence 156, Application US/09989730
: Publication No. US20020197674A1
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi J.
: APPLICANT: Baker, Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan L.
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Fong, Sherman
: APPLICANT: Gerder, Hanspeter
: APPLICANT: Gerlitsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Kijavlin, Ivar J.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2730PIC69
: CURRENT APPLICATION NUMBER: US/09/989,730
: PRIOR FILING DATE: 2001-11-20
: PRIOR APPLICATION NUMBER: 60/049787
: PRIOR FILING DATE: 1997-06-16
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/065186
: PRIOR FILING DATE: 1997-11-12
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066770
: PRIOR FILING DATE: 1997-11-24
: PRIOR APPLICATION NUMBER: 60/075945
: PRIOR FILING DATE: 1998-02-25
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/083322
: PRIOR FILING DATE: 1998-04-28
: PRIOR APPLICATION NUMBER: 60/084600
: PRIOR FILING DATE: 1998-05-07
: PRIOR APPLICATION NUMBER: 60/087106
: PRIOR FILING DATE: 1998-05-28
: PRIOR APPLICATION NUMBER: 60/087607
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087609
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087759
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087827
: PRIOR FILING DATE: 1998-06-03
: PRIOR APPLICATION NUMBER: 60/088021
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088025
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088026
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088028
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088029
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088030
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088033
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088326
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088167
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088202
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088212
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088217
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088655
: PRIOR FILING DATE: 1998-06-09
: PRIOR APPLICATION NUMBER: 60/088734
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088738
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088742
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088810
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088824
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088826
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088858
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/088861
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/088876
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/089105
: PRIOR FILING DATE: 1998-06-12
: PRIOR APPLICATION NUMBER: 60/089440
: PRIOR FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/089512
: PRIOR FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/089514
: PRIOR FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/089532
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089538
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089598
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089599
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089600
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089653
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089801
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/089907
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/089908
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/089947
: PRIOR FILING DATE: 1998-06-19
: PRIOR APPLICATION NUMBER: 60/089948
```

PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 62.8
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%

Length: 378
Matches: 8
Conservative: 0
Mismatch: 0

Query Match: 1.21% Indels: 0
DB: 9 Gaps: 0
US-09-513-151-3 (1-2041) x US-09-989-730-156 (1-378)
Oy 669 AGTCTCTAACCTTCATCCTT 692
Db 334 SerseleuthrleuAlaSerphe 341
RESULT 12
US-09-990-436-156
Sequence 156, Application US/09990436
Publication No. US20020196148A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tunas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C14
CURRENT APPLICATION NUMBER: US/09/990.436
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04

;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952

;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	62.8	Length:	378
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	9	Gaps:	0

US-09-513-151-3 (1-2041) x US-09-990-436-156 (1-378)
OY 669 AGTCTCTACCCCTTGATCTTT 692
|||||
Db 334 SerSerLeuThrLeuAlaSerPhe 341
RESULT 13
US-09-991-181-156
Sequence 156, Application US/09991181
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavir, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C53
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: US/09/991,181
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246

;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252

;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	62.8	Length:	378
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	9	Gaps:	0

US-09-513-151-3 (1-2041) x US-09-993-687-156 (1-378)

Oy 669 AGTCTCAACCTGATCCTT 692

Db 334 SerSertleuthleualaserpne 341
RESULT 15
US-09-989-734-156
Sequence 156, Application US/09989734
Publication No. US20030003531A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Geritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijaviri, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides, and Nucleic
FILE REFERENCE: P2730P1C64
CURRENT APPLICATION NUMBER: US/09/989,734
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088555
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254

PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	62.8	Length:	378
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	9	Gaps:	0

US-09-513-151-3 (1-2041) x US-09-989-734-156 (1-378)

OY 669 AGTCTCTAACCTTGATCCTTT 692
|||||
DB 334 SerSerLeuThrLeuAlaSerPhe 341

Search completed: April 21, 2003, 19:34:14
Job time : 41.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 21:27:17 : Search time 210 Seconds
(without alignments)
9763.385 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041
Sequence: 1 CTGCCATTAAGTCGCTCCG.....TTTACAGAGAAAAA 2041

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 639749 seqs, 502280978 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Published Applications.NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/ECT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/ECTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	473	23.2	717	US-09-764-853-166	Sequence 166, App
2	473	23.2	717	US-09-764-853-364	Sequence 364, App
3	318	15.6	318	US-09-796-692-4310	Sequence 4310, App
4	145	7.1	457	US-09-864-761-10050	Sequence 10050, A
5	138	6.8	138	US-09-864-761-26692	Sequence 26692, A
6	7	1.1	180557	US-10-003-806-6	Sequence 6, Appli
7	23	1.1	180557	US-10-003-806-9	Sequence 2777, Ap
8	22	1.1	1118	US-09-764-877-2777	Sequence 2770, Ap
9	22	1.1	1398	US-09-764-877-2780	Sequence 301, App
10	21	1.0	2154	US-09-822-846-301	Sequence 192, App
11	21	1.0	99916	US-09-910-943-192	Sequence 582, App
12	20	1.0	716	US-10-125-540-582	Sequence 582, App
13	20	1.0	1684	US-09-764-870-582	Sequence 857, App
14	20	1.0	1684	US-09-764-853-857	Sequence 1035, App
15	20	1.0	1684	US-09-764-860-1035	Sequence 214, App
16	20	1.0	1772	US-09-822-830A-214	Sequence 581, App
17	20	1.0	2396	US-10-125-540-581	Sequence 581, App
18	20	1.0	2396	US-09-764-870-581	Sequence 581, App
19	20	1.0	2396	US-09-764-870-581	Sequence 581, App

ALIGNMENTS

C	20	20	1.0	2396	10	US-09-764-853-856	Sequence 856, App
C	21	20	1.0	2396	10	US-09-764-860-1034	Sequence 1034, App
C	22	20	1.0	11216	9	US-09-764-872-490	Sequence 490, App
C	23	20	1.0	69327	10	US-09-777-921A-3	Sequence 3, Appli
C	24	20	1.0	198285	10	US-09-880-107-3814	Sequence 3814, App
C	25	20	1.0	465237	20	US-09-933-267A-1	Sequence 1, Appli
C	26	19	0.9	213	9	US-10-083-357-206	Sequence 206, App
C	27	19	0.9	384	10	US-09-983-965-5442	Sequence 5442, App
C	28	19	0.9	594	9	US-10-002-344A-59	Sequence 59, Appli
C	29	19	0.9	1179	10	US-09-770-445-52	Sequence 52, Appli
C	30	19	0.9	1659	9	US-09-852-797-24	Sequence 24, Appli
C	31	19	0.9	1659	10	US-09-853-161-24	Sequence 24, Appli
C	32	19	0.9	1659	10	US-09-852-659A-24	Sequence 24, Appli
C	33	19	0.9	2848	9	US-10-002-344A-60	Sequence 60, Appli
C	34	19	0.9	3564	9	US-10-091-504-1986	Sequence 1986, App
C	35	19	0.9	3564	10	US-09-764-869-1986	Sequence 1986, App
C	36	19	0.9	4548	9	US-10-091-504-2034	Sequence 2034, App
C	37	19	0.9	4548	10	US-09-764-869-2034	Sequence 2034, App
C	38	19	0.9	12566	10	US-10-091-504-2035	Sequence 2035, App
C	39	19	0.9	12566	10	US-09-764-869-2035	Sequence 2035, App
C	40	19	0.9	18000	10	US-09-800-631-17	Sequence 17, Appli
C	41	19	0.9	42000	9	US-10-081-563-25	Sequence 25, Appli
C	42	19	0.9	49136	10	US-09-768-877-1	Sequence 1, Appli
C	43	19	0.9	180216	10	US-09-835-232-6	Sequence 6, Appli
C	44	19	0.9	397658	10	US-09-813-320-3	Sequence 3, Appli
C	45	18	0.9	41	10	US-09-766-273-15	Sequence 15, Appli

RESULT 1
US-09-764-853-166
Sequence 166, Application US/09764853
Patent No. US20020090672A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P206
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 939
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 166
LENGTH: 717
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (619)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (625)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (655)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (673)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (713)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-853-166
Query Match 23.2% Score 473; DB 10; Length 717;
Best Local Similarity 99.7% Pred. No. 3.3e-239;
Matches 573; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
44 GTGGCAGTGGGCTCAGGGGCTGCAACGACCTTCTTCTAGTATTCGGGGCC 103
DB 44 GTGGCAGTGGGCTCAGGGGCTGCAACGACCTTCTTCTAGTATTCGGGGCC 103

Query Match	23.2%	Score 473	DB 10	Length 717
Best Local Similarity	99.7%	Pred. No. 3.3e-239		
Matches 573	Conservative	0	Mismatches 2	Indels 0
			Gaps	0

QY	44	GTGGGCAGTGGGCTTCAGGGGCTTCCAAACGACCCCTACCTCTTGTAGTATTCGCGGCC	103
Db	44	GTGGGCAATGGGCTTCAGGGGCTTCCAAACGACCCCTACCTCTTGTAGTATTCGCGGCC	103
QY	104	ACGGGACACGGGCAATCCACCGCTGGCTTGCAGCTAGGCCACGGCGTGGCGGTAGATC	163
Db	104	ACGGGACACGGGCAATCCACCGCTGGCTTGCAGCTAGGCCACGGCGTGGCGGTAGATC	163
QY	164	GTGAGCGGTGACTCCATGACAGGTATGAAAGGCTTACACATCATCACCACAAGTTTCT	223
Db	164	GTGAGCGGTGACTCCATGACAGGTATGAAAGGCTTACACATCATCACCACAAGTTTCT	223
QY	224	GCCCAAGACGAGAAATCTGCCGGCACCACATGATCAGCTTTTGGATCCTCTTGACCC	283
Db	224	GCCCAAGACGAGAAATCTGCCGGCACCACATGATCAGCTTTTGGATCCTCTTGACCC	283
QY	284	AATTACACAGTGGTGACCTTCAGAAATAGAGCAACTGCTCTGATTTAGATATATTTGCC	343
Db	284	AATTACACAGTGGTGACCTTCAGAAATAGAGCAACTGCTCTGATTTAGATATATTTGCC	343
QY	344	CGAAGCAAAATTTCTCTATTGTGTGGGAGAACCAATTTATCATTTGATCTCTGCTCG	403
Db	344	CGAAGCAAAATTTCTCTATTGTGTGGGAGAACCAATTTATCATTTGATCTCTGCTCG	403
QY	404	AAAGTCTTGTCAATTACCACGCCGAGAGATGGGACCTGGAAGGATTGACCGAAAA	463
Db	404	AAAGTCTTGTCAATTACCACGCCGAGAGATGGGACCTGGAAGGATTGACCGAAAA	463
QY	464	GTGAGCTTGAAGAGAGATGTTCTTGTACTTCACAACGCCCTAAGCCAGGTGACCCA	523
Db	464	GTGAGCTTGAAGAGAGATGTTCTTGTACTTCACAACGCCCTAAGCCAGGTGACCCA	523
QY	524	GAATGCGTGGCAGCTGCATCCACATGACACAAAGCGCCAGAGACTTGCAGATT	583
Db	524	GAATGCGTGGCAGCTGCATCCACATGACACAAAGCGCCAGAGACTTGCAGATT	583
QY	584	TTTGAGAAACAGGAATCTCTCATAGTGAATTTCT	618
Db	584	TTTGAGAAACAGGAATCTCTCATAGTGAATTTCT	618

RESULT 3

US-09-796-692-4310

US-09-796-692-4310, Application US/09796692

Publication No. US20020198362A1

GENERAL INFORMATION:

APPLICANT: Galiger, Alexander

APPLICANT: Algate, Paul A.

APPLICANT: Mannion, Jane

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND TREATMENT OF HEMATOLOGICAL MALIGNANCIES

FILE REFERENCE: 2077.001200

CURRENT APPLICATION NUMBER: US/09/796.692

CURRENT FILING DATE: 2001-03-01

PRIOR APPLICATION NUMBER: 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: 60/206,201

```
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4310
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-4310

Query Match
Best Local Similarity 100.0%; Score 318; DB 9; Length 318;
Pred. No. 2,3e-157;
Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1690 AAGAAGTTTATGTCCTGACTGCGCTAAATTAATTCAGATGCTTTGTAG 1749
    |||||||
DB 1 AAGAAGTTTATGTCCTGACTGCGCTAAATTAATTCAGATGCTTTGTAG 60

QY 1750 ATGACTGAGTATTTGTGACCAATTTGGAGTTCTAGATTTGATGCGAGAA 1809
    |||||||
DB 61 ATGACTGAGTATTTGTGACCAATTTGGAGTTCTAGATTTGATGCGAGAA 120

QY 1810 AGGCGCATCTCCATTGATGATTAAGTAAACCAACTAGTCTCGAATTTACAGAGA 1869
    |||||||
DB 121 AGGCGCATCTCCATTGATGATTAAGTAAACCAACTAGTCTCGAATTTACAGAGA 180

QY 1870 AGGAGGAATTCAGACTGAGGAAGCTGTGACATAGACTTGAACCAAGACTTTGAAT 1929
    |||||||
DB 181 AGGAGGAATTCAGACTGAGGAAGCTGTGACATAGACTTGAACCAAGACTTTGAAT 240

QY 1930 TTGCGAGCTGCTCATGTGAGATTATATCAGTCTGCTTTCTATTGACTCAATCT 1989
    |||||||
DB 241 TTGCGAGCTGCTCATGTGAGATTATATCAGTCTGCTTTCTATTGACTCAATCT 300

QY 1990 ATATTTTATTTAGATT 2007
    |||||||
DB 301 ATATTTTATTTAGATT 318

RESULT 4
US-09-864-761-10050/c
; Sequence 10050, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: CB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10050
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL033527.25
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
US-09-864-761-10050

Query Match
Best Local Similarity 7.1%; Score 145; DB 10; Length 457;
Pred. No. 5.3e-66;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 AGGCTATGAGGCTTAGCATCATCATCAGCAAGGTTTCTGCCAAGAGAGAGATCT 242
    |||||||
DB 450 AGGCTATGAGGCTTAGCATCATCATCAGCAAGGTTTCTGCCAAGAGAGAGATCT 391

QY 243 GCCGCGCACCATGATCAGCTTTGTGATCTCTTGTGACCAATTACAGAGTGTGACT 302
    |||||||
DB 390 GCCGCGCACCATGATCAGCTTTGTGATCTCTTGTGACCAATTACAGAGTGTGACT 331

QY 303 TCAGAAATAGAGCACTGCTCTGAT 327
    |||||||
DB 330 TCAGAAATAGAGCACTGCTCTGAT 306

RESULT 5
US-09-864-761-26692/c
; Sequence 26692, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
```


Query Match	1.18;	Score 23;	DB 12;	Length 180557
Best Local Similarity	100.0%;	Pred. No. 0.19;		
Matches 23;	Conservative 0;	Mismatches 0;	Indels 0	

RESULT 7
US-10-003-806-9/c

Query Match	1.1%;	Score 23;	DB 12;	Length 180557
Best Local Similarity	100.0%;	Pred. No. 0.19;		
Matches	23;	Conservative	0;	Mismatches 0
				Indels 0

RESULT 8
US-09-764-877-2777/c

```

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

```

;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005

CURRENT APPLICATION NUMBER: US/09/764,877

CURRENT FILING DATE: 2001-01-17

NUMBER OF SEO ID NOS: 4031

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2777
LENGTH: 1118
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-877-2777

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 1118;
Pred. No. 0.5;
Matches 22: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2019 AAATTTCACAGAAAAA 2040
|||
Db 221 AAATTTCACAGAAAAA 200

RESULT 9
US-09-764-877-2780/C

Sequence 2780, Application US/09764877
Patent No. US20020147140A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC005
CURRENT APPLICATION NUMBER: US/09/764,877
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 4031
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2780
LENGTH: 1398
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-877-2780

Query Match
Best Local Similarity 100.0%; Score 22; DB 10; Length 1398;
Pred. No. 0.5;
Matches 22: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2019 AAATTTCACAGAAAAA 2040
|||
Db 221 AAATTTCACAGAAAAA 200

RESULT 10
US-09-822-846-301/C

Sequence 301, Application US/09822846
Patent No. US20030027139A1
GENERAL INFORMATION:
APPLICANT: McCoy, John M.
APPLICANT: Lavallie, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Agostino, Michael J.
APPLICANT: Steininger II, Robert J.
APPLICANT: Bowman, Michael R.
APPLICANT: Spaulding, Vikki
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Fechtel, Kim
APPLICANT: Howes, Steven H.
APPLICANT: Resnick, Richard J.
APPLICANT: Gulukota, Kamalakara
APPLICANT: Graham, James R.
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
FILE REFERENCE: GIN 6400
CURRENT APPLICATION NUMBER: US/09/822,846
CURRENT FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/195,605
PRIOR FILING DATE: 2000-04-06

NUMBER OF SEQ ID NOS: 629
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 301
LENGTH: 2154
TYPE: DNA
ORGANISM: Homo sapiens
US-09-822-846-301

Query Match
Best Local Similarity 100.0%; Score 21; DB 9; Length 2154;
Pred. No. 1.7;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1566 CAGGAAGCATTTTTC 1586
|||
Db 1639 CAGGAAGCATTTTTC 1619

RESULT 11
US-09-816-095-3

Sequence 3, Application US/09816095
Patent No. US20020137164A1
GENERAL INFORMATION:
APPLICANT: Gan, Weiniu
TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
FILE REFERENCE: CL001147
CURRENT APPLICATION NUMBER: US/09/816,095
CURRENT FILING DATE: 2001-03-26
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 99916
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)..(99916)
OTHER INFORMATION: n - A,T,C or G
US-09-816-095-3

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 99916;
Pred. No. 2.1;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2019 AAATTTCACAGAAAAA 2039
|||
Db 43316 AAATTTCACAGAAAAA 43336

RESULT 12
US-09-910-943-192

Sequence 192, Application US/09910943
Patent No. US20020081610A1
GENERAL INFORMATION:
APPLICANT: Hemmati-Briandlou, Ali
TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
FILE REFERENCE: 7529/1G1480S1
CURRENT APPLICATION NUMBER: US/09/910,943
CURRENT FILING DATE: 2001-07-23
NUMBER OF SEQ ID NOS: 742
SOFTWARE: Patentln version 3.1
SEQ ID NO 192
LENGTH: 716
TYPE: DNA
ORGANISM: Xenopus laevis
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)..(716)
OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-192

Query Match 1.0%; Score 20; DB 10; Length 716;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2022 TTTACAGAGAAAAA 2041
DB 492 TTTACAGAGAAAAA 511

RESULT 13

US-10-125-540-582
; Sequence 582, Application US/10125540
; Publication No. US20030059675A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214CI
; CURRENT APPLICATION NUMBER: US/10/125,540
; CURRENT FILING DATE: 2002-04-19
; Prior Application removed - See File Wrapper or PALM
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 582
; LENGTH: 1684
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-125-540-582

Query Match 1.0%; Score 20; DB 9; Length 1684;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2019 AAATTACAGAAAAA 2038
DB 378 AAATTACAGAAAAA 397

RESULT 14

US-09-764-870-582
; Sequence 582, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 582
; LENGTH: 1684
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-870-582

Query Match 1.0%; Score 20; DB 10; Length 1684;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2019 AAATTACAGAAAAA 2038
DB 378 AAATTACAGAAAAA 397

RESULT 15

US-09-764-853-857/C
; Sequence 857, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206

; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 857
; LENGTH: 1684
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-853-857

Query Match 1.0%; Score 20; DB 10; Length 1684;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2019 AAATTACAGAAAAA 2038
DB 1307 AAATTACAGAAAAA 1288

Search completed: April 22, 2003, 00:12:53
Job time : 678 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 21:22:03 ; Search time 968 Seconds

(Without alignments)
9452.049 Million cell updates/sec

Title: US-09-513-151-3
Perfect score: 2041
Sequence: 1 CTCGCATTAACATGCGTCGCG.....TTTACAGCAAAAAA 2041

Scoring table: OLIGO_NUC
Gap 60.0 , Gapext 60.0

Searched: 5930009 seqs, 224146263 residues

Word size: 0

Total number of hits satisfying chosen parameters: 11860018

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

1: Pending-Patents_NA_New:*
2: /cgn2_6/p/ptodata/2/pna/US07_NEW.COMB.seq:*
3: /cgn2_6/p/ptodata/2/pna/US07_NEW.COMB.seq:*
4: /cgn2_6/p/ptodata/2/pna/US08_NEW.COMB.seq:*
5: /cgn2_6/p/ptodata/2/pna/US09_NEW.COMB.seq:*
6: /cgn2_6/p/ptodata/2/pna/US09_NEW.COMB.seq2:*
7: /cgn2_6/p/ptodata/2/pna/US09_NEW.COMB.seq3:*
8: /cgn2_6/p/ptodata/2/pna/US10_NEW.COMB.seq:*
9: /cgn2_6/p/ptodata/2/pna/US10_NEW.COMB.seq2:*
10: /cgn2_6/p/ptodata/2/pna/US60_NEW.COMB.seq:*
11: /cgn2_6/p/ptodata/2/pna/US60_NEW.COMB.seq2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	896	43.9	2130	US-09-659-671A-134	Sequence 134, App
2	896	43.9	2130	US-10-380-731-119	Sequence 119, App
3	767	37.6	1271	US-10-170-235-41124	Sequence 41124, A
4	765	37.5	1719	US-10-170-235-41123	Sequence 41123, A
5	765	37.5	2119	US-10-170-235-40554	Sequence 40554, A
6	751	36.8	1749	US-10-342-887-1692	Sequence 1692, A
7	622	30.5	1271	US-60-453-135-276	Sequence 276, App
8	622	30.5	1271	US-60-453-050-276	Sequence 276, App
9	622	30.5	1719	US-60-453-135-274	Sequence 274, App
10	622	30.5	1719	US-60-453-050-274	Sequence 274, App
11	622	30.5	2119	US-60-453-135-275	Sequence 275, App
12	622	30.5	2119	US-60-453-050-275	Sequence 275, App
13	404	19.8	555	US-09-532-315B-3472	Sequence 3472, App
14	269	13.2	392	US-09-532-315B-3471	Sequence 3471, App
15	265	13.0	293	US-09-532-315B-9521	Sequence 9521, App
16	264	12.9	264	US-09-532-315B-9514	Sequence 9514, App
17	256	12.5	256	US-09-532-315B-9508	Sequence 9508, App
18	255	12.5	255	US-09-532-315B-9511	Sequence 9511, App
19	247	12.1	247	US-09-532-315B-9505	Sequence 9505, App
20	237	11.6	266	US-09-532-315B-9507	Sequence 9507, App
21	230	11.3	289	US-09-532-315B-9519	Sequence 9519, App
22	193	9.5	240	US-09-532-315B-9515	Sequence 9515, App

23	192	9.4	265	US-09-532-315B-9518	Sequence 9518, App
24	189	9.3	189	US-09-532-315B-9500	Sequence 9500, App
25	182	8.9	245	US-09-532-315B-9523	Sequence 9523, App
26	180	8.8	263	US-09-532-315B-9510	Sequence 9510, App
27	179	8.8	180	US-09-532-315B-9503	Sequence 9503, App
28	174	8.5	334	US-09-912-293-109604	Sequence 109604, App
29	172	8.4	232	US-09-532-315B-9529	Sequence 9529, App
30	150	7.3	201	US-60-453-135-18354	Sequence 18354, A
31	150	7.3	201	US-60-453-135-18357	Sequence 18357, A
32	150	7.3	201	US-60-453-135-18362	Sequence 18362, A
33	150	7.3	201	US-60-453-135-18364	Sequence 18364, A
34	150	7.3	201	US-60-453-135-18368	Sequence 18368, A
35	150	7.3	201	US-60-453-050-18354	Sequence 18354, A
36	150	7.3	201	US-60-453-050-18357	Sequence 18357, A
37	150	7.3	201	US-60-453-050-18362	Sequence 18362, A
38	150	7.3	201	US-60-453-050-18364	Sequence 18364, A
39	150	7.3	201	US-60-453-050-18368	Sequence 18368, A
40	149	7.3	255	US-09-532-315B-9517	Sequence 9517, App
41	148	7.3	255	US-09-532-315B-9520	Sequence 9520, App
42	145	7.1	457	US-10-203-138A-1534	Sequence 1534, App
43	144	7.1	262	US-09-532-315B-9528	Sequence 9528, App
44	139	6.8	500	US-09-912-293-111181	Sequence 111181, App
45	138	6.8	138	US-10-203-138A-6661	Sequence 6661, App

ALIGNMENTS

```
RESULT 1
US-09-659-671A-134
Sequence 134, Application US/09659671A
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenshua
APPLICANT: Zhou, Ping
APPLICANT: Asundi, Vinod
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyang J.
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Weinman, Tom
APPLICANT: Drmanac, Radolje T.
TITLE OF INVENTION: Novel Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 794
CURRENT APPLICATION NUMBER: US/09/659, 671A
CURRENT FILING DATE: 2000-09-11
NUMBER OF SEQ ID NOS: 475
SOFTWARE: pf-fl-genes Version 2.0
SEQ ID NO 134
LENGTH: 2130
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (191)..(1426)
US-09-659-671A-134
Query Match 43.9%; Score 896; DB 6; Length 2130;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 896; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

44 GTGGCAGTGGCTAGGGCCCTGACAGGACCTCTTGTAGTATTTCGGGCC 103
DB GTGGCAGTGGCTAGGGCCCTGACAGGACCTCTTGTAGTATTTCGGGCC 115
56 GTGGCAGTGGCTAGGGCCCTGACAGGACCTCTTGTAGTATTTCGGGCC 115
QY ACGGCGACCGGCAATTCACGCTGCTAGGCGACGCGCGCGTACATC 163
DB ACGGCGACCGGCAATTCACGCTGCTAGGCGACGCGCGCGTACATC 175
QY 164 GTGACGCTGACTCCATGCTATGAGAGGCTATGACCAACAGAGTTCT 223

```
|||||
Db 176 GTGAGCCCTGACTGCATGCAAGTCTATGAGGCTTACATCATCACCACCAAGTTCT 235
OY 224 GCCCAGAGCAGAGAAATCTCCGGCACACATGATGAGCTTTGTGATCTTTGTGACC 283
Db 236 GCCCAAGAGCAGAAATCTCCGGCACACATGATGAGCTTTGTGATCTTTGTGACC 295
OY 284 AATTACACAGTGGTGCCTCAGAAAATAGCAAACTGCTGATTTGAAGATATTTGGC 343
Db 296 AATTACACAGTGGTGCCTCAGAAAATAGCAAACTGCTGATTTGAAGATATTTGGC 355
OY 344 CGAGACAAATTCCTATTGTTGGAGAGAACCAATTATTACATTGAATCTCTGCG 403
Db 356 CGAGACAAATTCCTATTGTTGGAGAGAACCAATTATTACATTGAATCTCTGCG 415
OY 404 AAGTTCTTGTCAATACCAAGCCCGCAGAGATGGGACCTGAGAAATGATTGACGAAA 463
Db 416 AAGTTCTTGTCAATACCAAGCCCGCAGAGATGGGACCTGAGAAATGATTGACGAAA 475
OY 464 GTGAGCTTGAAGAGAGATGGTCTGACTTCCACAAGGCTTAACGCGAGTGAGCCCA 523
Db 476 GTGAGCTTGAAGAGAGATGGTCTGACTTCCACAAGGCTTAACGCGAGTGAGCCCA 535
OY 524 GAAATGGCTGCCAAGCTGCATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGATT 583
Db 536 GAAATGGCTGCCAAGCTGCATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGATT 595
OY 584 TTGGAAGAAACAGCAATCTCTCATAGTGAATTTCTCATGTCACATACGGAAGAAGT 643
Db 596 TTGGAAGAAACAGCAATCTCTCATAGTGAATTTCTCATGTCACATACGGAAGAAGT 655
OY 644 GGTGGTCCCTTGAGGTCCTCTGAAGTCTCTCAACCCCTTGACCTTTGGCTTCATGCT 703
Db 656 GGTGGTCCCTTGAGGTCCTCTGAAGTCTCTCAACCCCTTGACCTTTGGCTTCATGCT 715
OY 704 GACCAGCAGTCTAGATGAGCGCTTGATAGAGGCTGATGACATGCTTGTGCTGGG 763
Db 716 GACCAGCAGTCTAGATGAGCGCTTGATAGAGGCTGATGACATGCTTGTGCTGGG 775
OY 764 CTCTTGAGGAGACTAGAGATTTTTCACAGACGCTATATATAGAGAAATGTTTGGAAAT 823
Db 776 CTCTTGAGGAGACTAGAGATTTTTCACAGACGCTATATATAGAGAAATGTTTGGAAAT 835
OY 824 AGCAGAGACTATCAACATGATCTTCCATCAATGAGCTTCAAGGAATTTCCAGATAC 883
Db 836 AGCAGAGACTATCAACATGATCTTCCATCAATGAGCTTCAAGGAATTTCCAGATAC 895
OY 884 CTGATCAGTGAAGGAAATGCACTGTGAGACTAGTAACCAAGCTTCTAAGAAAG 939
Db 896 CTGATCAGTGAAGGAAATGCACTGTGAGACTAGTAACCAAGCTTCTAAGAAAG 951

RESULT 2
US-10-380-731-119
; Sequence 119, Application US/10380731
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-114
; CURRENT APPLICATION NUMBER: US/10/380,731
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 09/659,671
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 888
; SOFTWARE: Custom
; SEQ ID NO 119
; LENGTH: 2130
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (191)..(1426)
US-10-380-731-119
```

```
Query Match 43.9%; Score 896; DB 8; Length 2130;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 896; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 44 GTGGGAGTGGGCTCAGGGGCTGCAACGAGACCTTACTTGTAGTATCTCGGGGCC 103
Db 56 GTGGGAGTGGGCTCAGGGGCTGCAACGAGACCTTACTTGTAGTATCTCGGGGCC 115
OY 104 ACGGGACCGGCAATTCACAGCTGCGCTTGCAGCTAGGCGGCGCTCGGGGTGAGATC 163
Db 116 ACGGGACCGGCAATTCACAGCTGCGCTTGCAGCTAGGCGGCGCTCGGGGTGAGATC 175
OY 164 GTAGGCTACTTCATGACAGGTCTATGAAGGCTTAGACATCATCAACAAAGTTTCT 223
Db 176 GTAGGCTACTTCATGACAGGTCTATGAAGGCTTAGACATCATCAACAAAGTTTCT 235
OY 224 GCCCAAGAGCAGAAATCTCCGGCACACATGATGAGCTTTGTGATCTTTGTGACC 283
Db 236 GCCCAAGAGCAGAAATCTCCGGCACACATGATGAGCTTTGTGATCTTTGTGACC 295
OY 284 AATTACACAGTGGTGCCTCAGAAAATAGCAAACTGCTGATTTGAAGATATTTGGC 343
Db 296 AATTACACAGTGGTGCCTCAGAAAATAGCAAACTGCTGATTTGAAGATATTTGGC 355
OY 344 CGAGACAAATTCCTATTGTTGGAGAGAACCAATTATTACATTGAATCTCTGCTGCG 403
Db 356 CGAGACAAATTCCTATTGTTGGAGAGAACCAATTATTACATTGAATCTCTGCTGCG 415
OY 404 AAGTTCTTGTCAATACCAAGCCCGCAGAGATGGGACCTGAGAAATGATTGACGAAA 463
Db 416 AAGTTCTTGTCAATACCAAGCCCGCAGAGATGGGACCTGAGAAATGATTGACGAAA 475
OY 464 GTGAGCTTGAAGAGAGATGGTCTGACTTCCACAAGGCTTAACGCGAGTGAGCCCA 523
Db 476 GTGAGCTTGAAGAGAGATGGTCTGACTTCCACAAGGCTTAACGCGAGTGAGCCCA 535
OY 524 GAAATGGCTGCCAAGCTGCATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGATT 583
Db 536 GAAATGGCTGCCAAGCTGCATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGATT 595
OY 584 TTGGAAGAAACAGCAATCTCTCATAGTGAATTTCTCATGTCACATACGGAAGAAGT 643
Db 596 TTGGAAGAAACAGCAATCTCTCATAGTGAATTTCTCATGTCACATACGGAAGAAGT 655
OY 644 GGTGGTCCCTTGAGGTCCTCTGAAGTCTCTCAACCCCTTGACCTTTGGCTTCATGCT 703
Db 656 GGTGGTCCCTTGAGGTCCTCTGAAGTCTCTCAACCCCTTGACCTTTGGCTTCATGCT 715
OY 704 GACCAGCAGTCTAGATGAGCGCTTGATAGAGGCTGATGACATGCTTGTGCTGGG 763
Db 716 GACCAGCAGTCTAGATGAGCGCTTGATAGAGGCTGATGACATGCTTGTGCTGGG 775
OY 764 CTCTTGAGGAGACTAGAGATTTTTCACAGACGCTATATATAGAGAAATGTTTGGAAAT 823
Db 776 CTCTTGAGGAGACTAGAGATTTTTCACAGACGCTATATATAGAGAAATGTTTGGAAAT 835
OY 824 AGCAGAGACTATCAACATGATCTTCCATCAATGAGCTTCAAGGAATTTCCAGATAC 883
Db 836 AGCAGAGACTATCAACATGATCTTCCATCAATGAGCTTCAAGGAATTTCCAGATAC 895
OY 884 CTGATCAGTGAAGGAAATGCACTGTGAGACTAGTAACCAAGCTTCTAAGAAAG 939
Db 896 CTGATCAGTGAAGGAAATGCACTGTGAGACTAGTAACCAAGCTTCTAAGAAAG 951

RESULT 3
US-10-170-235-41124
; Sequence 41124, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREO
```

```
FILE REFERENCE: CL001380
CURRENT APPLICATION NUMBER: US/10/170.235
CURRENT FILING DATE: 2003-03-17
NUMBER OF SEQ ID NOS: 42514
SEQ ID NO 41124
LENGTH: 1271
TYPE: DNA
ORGANISM: HUMAN
US-10-170-235-41124

Query Match      37.6%; Score 767; DB 8; Length 1271;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1087; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

OY 937 AGACCTGTCCTCCATGTCCTCCCTGCTGTATGCTTACAGATATGATGTCTGCAAGT 996
    |||||
DB 181 AGGACCTGTCCTCCATGTCCTCCCTGCTGTATGCTTACAGATATGATGTCTGCAAGT 240
    |||||

OY 997 GGAGAGTCTGTCTTGAACCTGCTTGAACCTGCTTGAACCTGCTTGAACCTGCTTGAACCT 1056
    |||||
DB 241 GGAGAGTCTGTCTTGAACCTGCTTGAACCTGCTTGAACCTGCTTGAACCTGCTTGAACCT 300
    |||||

OY 1057 GCCTACAGCCACTCCCAATTAAGATGCCATACATGAGCTGAGAACAGAGAAGTTATCA 1116
    |||||
DB 301 GCCTACAGCCACTCCCAATTAAGATGCCATACATGAGCTGAGAACAGAGAAGTTATCA 360
    |||||

OY 1117 CCTGTGTACCTCTGTATGATCATCATTTGGGATGCCAATGGCCAGCCACATATA 1176
    |||||
DB 361 CCTGTGTACCTCTGTATGATCATCATTTGGGATGCCAATGGCCAGCCACATATA 420
    |||||

OY 1177 ATCCAAATCCCACTGAGCAACCTGAGAAAGAAAGATGGACTGAGTGTGTCAA 1236
    |||||
DB 421 ATCCAAATCCCACTGAGCAACCTGAGAAAGAAAGATGGACTGAGTGTGTCAA 480
    |||||

OY 1237 CACCATAGAAAGTCAGAGTGTTCCTCCAGACTATTAACAAGAACCTAAAGGAGGATC 1296
    |||||
DB 481 CACCATAGAAAGTCAGAGTGTTCCTCCAGACTATTAACAAGAACCTAAAGGAGGATC 540
    |||||

OY 1297 CCCAGAGGAGATGATCAAGAGCTGAAGCTTTTAAAGACATGTCCAGTGGCCTT 1356
    |||||
DB 541 CCCAGAGGAGATGATGATCAAGAGCTGAAGCTTTTAAAGACATGTCCAGTGGCCTT 600
    |||||

OY 1357 TGGAAAGTGTGGGGATGTCAGATGAGAGGAGGATGTTTGTCTCTCCAGTGGGC 1416
    |||||
DB 601 TGGAAAGTGTGGGGATGTCAGATGAGAGGAGGATGTTTGTCTCTCCAGTGGGC 660
    |||||

OY 1417 AAAGAGTGTATGCGGAATCTGTGATAGCAGAAAGCTCCACCATTTTCTTTGAT 1476
    |||||
DB 661 AAAGAGTGTATGCGGAATCTGTGATAGCAGAAAGCTCCACCATTTTCTTTGAT 720
    |||||

OY 1477 GTGCTTTAAAGTCTCAGCTTCTATATATAGAAACAGAGCTTTGTGACGTCCTTGTG 1536
    |||||
DB 721 GTGCTTTAAAGTCTCAGCTTCTATATATAGAAACAGAGCTTTGTGACGTCCTTGTG 780
    |||||

OY 1537 TGGCTGATGTGTCGGAATGATGATGATGAGAAAGCTTTTCTTTCTTTGAACCTT 1596
    |||||
DB 781 TGGCTGATGTGTCGGAATGATGATGATGAGAAAGCTTTTCTTTCTTTGAACCTT 840
    |||||

OY 1597 AAAGTCTTATTTAAAGAGCAGACAGATTCACATTTTATACATGAGATCTTCTTT 1656
    |||||
DB 841 AAAGTCTTATTTAAAGAGCAGACAGATTCACATTTTATACATGAGATCTTCTTT 900
    |||||

OY 1657 GTGCTGATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1716
    |||||
DB 901 GTGCTGATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 959
    |||||

OY 1717 CTAATATATCTAATTTTCCAGATGCTTTTGTGATGATGATGATGATGATGATGATGATG 1776
    |||||
DB 960 CTAATATATCTAATTTTCCAGATGCTTTTGTGATGATGATGATGATGATGATGATGATG 1019
    |||||

OY 1777 TTGGAGATCTAGATTTAGTGAATGAGAGAAAGGCCATCTCCATTTGAGATGATTAAG 1836
    |||||
DB 1020 TTGGAGATCTAGATTTAGTGAATGAGAGAAAGGCCATCTCCATTTGAGATGATTAAG 1079
    |||||
```

```
OY 1837 TGAACCAACTAGTCTCTCGAATCTCTACAGAGAGGAGGATCATGAGTGAAGAGCTGT 1896
    |||||
DB 1080 TGAACCAACTAGTCTCTCGAATCTCTACAGAGAGGAGGATCATGAGTGAAGAGCTGT 1139
    |||||

OY 1897 GACATAGACTTGAAGACCAAGACTTTGAATTTGGCAGCTGCTCATGTGTGACTATTT 1956
    |||||
DB 1140 GACATAGACTTGAAGACCAAGACTTTGAATTTGGCAGCTGCTCATGTGTGACTATTT 1199
    |||||

OY 1957 ATCACTGCTGTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2016
    |||||
DB 1200 ATCACTGCTGTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1259
    |||||

OY 2017 AAAAATTTACAA 2028
    |||||
DB 1260 AAAAATTTACAA 1271
    |||||

RESULT 4
US-10-170-235-41123
Sequence 41123, Application US/10170235
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig
TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF
TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THERE
FILE REFERENCE: CL001380
CURRENT APPLICATION NUMBER: US/10/170.235
CURRENT FILING DATE: 2003-03-17
NUMBER OF SEQ ID NOS: 42514
SEQ ID NO 41123
LENGTH: 1719
TYPE: DNA
ORGANISM: HUMAN
US-10-170-235-41123

Query Match      37.5%; Score 765; DB 8; Length 1719;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1085; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

OY 938 GACCTGTGCTCCATTTGTCCTCCCTGCTGTATGCTTACAGTATCTGATGTCTGAGTGG 998
    |||||
DB 631 GACCTGTGCTCCATTTGTCCTCCCTGCTGTATGCTTACAGTATCTGATGTCTGAGTGG 690
    |||||

OY 999 AGAGTGTGCTTCTTGAACCTGCTTTGAATGCTGCAAAAGTTTATCCAGGCGCACAAAGC 1058
    |||||
DB 691 AGAGTGTGCTTCTTGAACCTGCTTTGAATGCTGCAAAAGTTTATCCAGGCGCACAAAGC 750
    |||||

OY 1059 CTACAGCCACTCCCAATTAAGATGCTATACATGAAGCTGAGAACAGAGATATACCC 1118
    |||||
DB 751 CTACAGCCACTCCCAATTAAGATGCTATACATGAAGCTGAGAACAGAGATATACCC 810
    |||||

OY 1119 TGTGTGACCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1178
    |||||
DB 811 TGTGTGACCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 870
    |||||

OY 1179 CCAATGCCACTTGAACCACTGAGAAAGAAAGAAAGATTTGAGTCCAGATGCTGTCAACA 1238
    |||||
DB 871 CCAATGCCACTTGAACCACTGAGAAAGAAAGAAAGATTTGAGTCCAGATGCTGTCAACA 930
    |||||

OY 1239 CCATGAAAGTGAAGTGTTCCTCCAGACTATTAACAAGACCTAAAGGAGGATGCC 1298
    |||||
DB 931 CCATGAAAGTGAAGTGTTCCTCCAGACTATTAACAAGACCTAAAGGAGGATGCC 990
    |||||

OY 1299 CAGGCGAAGATATCAAGAGCTGGAATGAGAGCTTTAAGAGACATGCTCCAGTGGCTTTG 1358
    |||||
DB 991 CAGGCGAAGATATCAAGAGCTGGAATGAGAGCTTTAAGAGACATGCTCCAGTGGCTTTG 1050
    |||||

OY 1359 GAAAGTGTGGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1418
    |||||
DB 1051 GAAAGTGTGGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1110
    |||||

OY 1419 AGAGTGTATGCGGAATTTCTGTGATAGCAGAAAGCTCCACCATTTTCTTTGATGT 1478
    |||||
```



```

: APPLICANT: He, Yudong
: APPLICANT: Linsley, Peter S.
: APPLICANT: Mao, Mao
: APPLICANT: Roberts, Christopher J.
: APPLICANT: Van 't Veer, Laura Johanna
: APPLICANT: Van de Vijver, Marc J.
: APPLICANT: Bernards, Rene
: TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
: FILE REFERENCE: 9301-188-999
: CURRENT APPLICATION NUMBER: US/10/342,887
: CURRENT FILING DATE: 2003-01-15
: PRIOR APPLICATION NUMBER: 60/298,918
: PRIOR FILING DATE: 2001-06-18
: PRIOR APPLICATION NUMBER: 60/380,710
: PRIOR FILING DATE: 2002-05-14
: PRIOR APPLICATION NUMBER: 10/172,118
: PRIOR FILING DATE: 2002-06-14
: NUMBER OF SEQ ID NOS: 2699
: SEQ ID NO: 1692
: LENGTH: 1749
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-342-887-1692

```

Query Match 36.8%; Score 751; DB 9; Length 1749;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1021: Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

Qy 1001 GAGTCTGTTCTGAACCTGCTCTTGAATGTCGCAAAAGTTTCATCCAGGGCCACAGACCT 1060
Db 706 GAGTCTGTTCTGAACCTGCTCTTGAATGTCGCAAAAGTTTCATCCAGGGCCACAGACCT 765
Qy 1061 AAGCCACCTCAATAAGATGGCCATTCATGAAAGTCGAGAAACAGAGAGTATATCCCTG 1120
Db 766 AAGCCACCTCAATAAGATGGCCATTCATGAAAGTCGAGAAACAGAGAGTATATCCCTG 825
Qy 1121 TGTGACCTCTGATGATCATCATGATGGGATCCGATGGGACGACATATAATCC 1180
Db 826 TGTGACCTCTGATGATCATCATGATGGGATCCGATGGGACGACATATAATCC 885
Qy 1181 AATCCACTTGAACCACTGAAGAAAGAGAGATTGGACTCAGATCCTGTCAACACC 1240
Db 886 AATCCACTTGAACCACTGAAGAAAGAGAGATTGGACTCAGATCCTGTCAACACC 945
Qy 1241 ATAGAAAGTCAGAGTGTTCCTCCAGCTATACAAAGAACCTTAAGGGAAGGATCCCA 1300
Db 946 ATAGAAAGTCAGAGTGTTCCTCCAGCTATACAAAGAACCTTAAGGGAAGGATCCCA 1005
Qy 1301 GGACAGATGATCAAGAGCTGAAGATGCAGCTTTAAGAGACATGTCCAGTGGCTTTGA 1360
Db 1006 GGACAGATGATCAAGAGCTGAAGATGCAGCTTTAAGAGACATGTCCAGTGGCTTTGA 1065
Qy 1361 AAGGTGGTGGGATCCAGTTCAAGAGAGGAGGATGTTGTCTCCAGTCTGGGCAAG 1420
Db 1066 AAGGTGGTGGGATCCAGTTCAAGAGAGGAGGATGTTGTCTCCAGTCTGGGCAAG 1125
Qy 1421 GAGTGTATGCGAATTTCTGATAGCAGAAAGCTCCACCATTTCTTTGATGTGG 1480
Db 1126 GAGTGTATGCGAATTTCTGATAGCAGAAAGCTCCACCATTTCTTTGATGTGG 1185
Qy 1481 TTTTAAAGTCAAGTCTCTATATATGAAGACAGAGCTGTGTCAGCTCCTGTGTGGC 1540
Db 1186 TTTTAAAGTCAAGTCTCTATATATGAAGACAGAGCTGTGTCAGCTCCTGTGTGGC 1245
Qy 1541 TGAATGCTGCAAAATGATGATGATGAGGAAAGCATTTTCTTTTGAACCTTAAG 1600
Db 1246 TGAATGCTGCAAAATGATGATGATGAGGAAAGCATTTTCTTTTGAACCTTAAG 1305
Qy 1601 GTTCTATATTAAGACACAGATTCACATTTTATATACATGAGATCTTCTTTGTGG 1660
Db 1306 GTTCTATATTAAGACACAGATTCACATTTTATATACATGAGATCTTCTTTGTGG 1365
Qy 1661 TGAATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1720

```

```

Db 1366 TGAATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1424
Qy 1721 AATATCTAATTTCCAGATGCTTTGTGATGATGATGATGATGATGATGATGATGATGAT 1780
Db 1425 AATATCTAATTTCCAGATGCTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1484
Qy 1781 GAGTCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1840
Db 1485 GAGTCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1544
Qy 1841 CCAAACTAGTTCTCGAATTTCAAGAGAGAGGAGGAAATCAGACTGAGGAACTGTGACA 1900
Db 1545 CCAAACTAGTTCTCGAATTTCAAGAGAGAGGAGGAAATCAGACTGAGGAACTGTGACA 1604
Qy 1901 TAGGACTTGAACCAAGACCTTGAATTTTCCAGCTGCTCATGTGATGATTTATATCA 1960
Db 1605 TAGGACTTGAACCAAGACCTTGAATTTTCCAGCTGCTCATGTGATGATTTATATCA 1664
Qy 1961 CTGCTGCTTTCTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2020
Db 1665 CTGCTGCTTTCTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1724
Qy 2021 ATTTA 2025
Db 1725 ATTTA 1729

```

RESULT 7
US-60-453-135-276

; Sequence 276, Application US/60453135

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; APPLICANT: IAKOUBOVA, Olga

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; FILE REFERENCE: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THERE

; CURRENT APPLICATION NUMBER: US/60/453,135

; CURRENT FILING DATE: 2003-03-10

; NUMBER OF SEQ ID NOS: 82762

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO: 276

; LENGTH: 1271

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-60-453-135-276

Query Match 30.5%; Score 622; DB 11; Length 1271;

Best Local Similarity 99.4%; Pred. No. 1; 4e-298;

Matches 992: Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```

Qy 1031 GTGCAAGTTTCATCCAGGCGCAGCAAGCTTACAGCCACTGCATTAAGATGCCATTCAT 1090
Db 275 GTGCAAGTTTCATCCAGGCGCAGCAAGCTTACAGCCACTGCATTAAGATGCCATTCAT 334
Qy 1091 GAAAGTGAAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1150
Db 335 GAAAGTGAAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 394
Qy 1151 GATCGGGAATGGGACGACATTAATTAATCCCACTTGAACCACTGGAAGAAAGA 1210
Db 395 GATCGGGAATGGGACGACATTAATTAATCCCACTTGAACCACTGGAAGAAAGA 454
Qy 1211 AGAAGATGAGTCAAGTCTGTCAACACCATTAAGAAAGTCAAGTCTGTCCAGACTAT 1270
Db 455 AGAAGATGAGTCAAGTCTGTCAACACCATTAAGAAAGTCAAGTCTGTCCAGACTAT 514
Qy 1271 AACCAAGACCTTAAGAGAGAGATGCCAGGCGCAAGATGATGATGATGATGATGATGAT 1330
Db 515 AACCAAGACCTTAAGAGAGAGATGCCAGGCGCAAGATGATGATGATGATGATGATGAT 574
Qy 1331 GTTAAAGACATGTCAGATGCTTTGGAAGAGTGGGATCCAGTTCAGAGAGAG 1390

```


	Query Match	30.5%	Score 622	DB 11	Length 1271
	Best Local Similarity	99.4%	Pred. No. 1	4e-298	
	Matches 992	Conservative 0	Mismatches 5	Indels 1	Gaps 1
Oy	1031	GTGCAAGTTCATCTCAGGGCCACAGGCTCTACAGCCACTCCATATAGATGCCATACAAAT	1090		
Ob	275	GTGCAAGTTCATCTCAGGGCCACAGGCTCTACAGCCACTCCATATAGATGCCATACAAAT	334		

RESULT 9
 US-60-453-135-274
 Sequence 274, Application US/60453135
 GENERAL INFORMATION:
 APPLICANT: CARCILL, Michele
 APPLICANT: IAKOUBOVA, Olga
 TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

Db 1263 AACGATTTTCTTTGAACTTAAGTCTATATTAAGACAGACAGATTCCA 1322
Qy 1631 CATTTTATACATGAGATCTCTTTGTGTGAATACCAGATTGCACTCCCTTAA 1690
Db 1323 CATTTTATACATGAGATCTCTTTGTGTGAATACCAGATTGCACTCCCTTAA 1381
Qy 1691 AAGAGTTTATGTCCTGAGCTGGCTAAATATCTAATTTCCAGATGCTTTGTAGA 1750
Db 1382 AAGAGTTTATGTCCTGAGCTGGCTAAATATCTAATTTCCAGATGCTTTGTAGA 1441
Qy 1751 TGACTGAATATTTGTAGCCACATATTTGGAGTTCTAGATTGAGTAATGGCAGAAA 1810
Db 1442 TGACTGAATATTTGTAGCCACATATTTGGAGTTCTAGATTGAGTAATGGCAGAAA 1501
Qy 1811 GGGCATTCTCATTTGAGATGATTAAGTAAGCAACCAACTAGTTCTCGAATTTACAGAAA 1870
Db 1502 GGGCATTCTCATTTGAGATGATTAAGTAAGCAACCAACTAGTTCTCGAATTTACAGAAA 1561
Qy 1871 GAGGGAATCAGACTGAGAAAGCTGTACATAGACTTGAAGACCAAGACTTTGAATT 1930
Db 1562 GAGGGAATCAGACTGAGAAAGCTGTACATAGACTTGAAGACCAAGACTTTGAATT 1621
Qy 1931 TGGCAGCTGCATGTGTGAGTTATATCACTGCTCTTTCTTCTATGAGTTACAATCTA 1990
Db 1822 TGGCAGCTGCATGTGTGAGTTATATCACTGCTCTTTCTTCTATGAGTTACAATCTA 1681
Qy 1991 TATTTTATGAACTTAAATAAAGAAAATTTACAA 2028
Db 1682 TATTTTATGAACTTAAATAAAGAAAATTTACAA 1719

RESULT 11

US-60-453-135-275
; Sequence 275, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453.135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 275
; LENGTH: 2119
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-275

Query Match 30.5%; Score 622; DB 11; Length 2119;
Best Local Similarity 99.4%; Pred. No. 1,4e-298;
Matches 992; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
Qy 1031 GTGCAAGTTTCATCCAGGGGACACAGCTCAACAGCACTCCCAATAAAGATGCCATACAT 1090
Db 1123 GTGCAAGTTTCATCCAGGGGACACAGCTCAACAGCACTCCCAATAAAGATGCCATACAT 1182
Qy 1091 GAACTGAGAAAGAGAGAGATTATACCTGTGTGACCTCTGTGATGCAATCATATTGGG 1150
Db 1183 GAACTGAGAAAGAGAGAGATTATACCTGTGTGACCTCTGTGATGCAATCATATTGGG 1242
Qy 1151 GATGCGCAATGGCAGCGCACATATAATCCAAATCCCACTTGAACCAACTGAGAAAAGA 1210
Db 1243 GATGCGCAATGGCAGCGCACATATAATCCAAATCCCACTTGAACCAACTGAGAAAAGA 1302
Qy 1211 AGAGATTGGACTGAGTGTGTCACACATAGAAAGTGAAGTGTTCCTCCGCACTAT 1270
Db 1303 AGAGATTGGACTGAGTGTGTCACACATAGAAAGTGAAGTGTTCCTCCGCACTAT 1362
Qy 1271 AACAAAGAACCTAAAGGAGAGGATCCCGAGGCGAGATGATCAAGAGCTGAATGCGAGC 1330
Db 1363 AACAAAGAACCTAAAGGAGAGGATCCCGAGGCGAGATGATCAAGAGCTGAATGCGAGC 1422

Qy 1331 GTTTAAGAGACATGTCAGTGGCTTTTGAAAGTGTGGGGATCCAGTCCAGAGGAG 1390
Db 1423 GTTTAAGAGACATGTCAGTGGCTTTTGAAAGTGTGGGGATCCAGTCCAGAGGAG 1482
Qy 1391 GGGATGTTGTGTCCTCCAGTCTGGGCAAGAGAGTGTATGCGGAATTTCTGCTATACAG 1450
Db 1483 GGGATGTTGTGTCCTCCAGTCTGGGCAAGAGAGTGTATGCGGAATTTCTGCTATACAG 1542
Qy 1451 AAAAGCTCCACATTTCTTTTGATGTGTTTTAAAGTCTCACGTTCTCTATATAGAA 1510
Db 1543 AAAAGCTCCACATTTCTTTTGATGTGTTTTAAAGTCTCACGTTCTCTATATAGAA 1602
Qy 1511 ACAGAGTCTTGTGACACTCTTGTGTGCGTGTATGCTGTGGAATATGATGATTCAGGA 1570
Db 1603 ACAGAGTCTTGTGACACTCTTGTGTGCGTGTATGCTGTGGAATATGATGATTCAGGA 1662
Qy 1571 AAGCATTTTCTTTCTTTGACCTTTAAGTTTATATTAAGACAGACAGATTTCA 1630
Db 1663 AAGCATTTTCTTTCTTTGACCTTTAAGTTTATATTAAGACAGACAGATTTCA 1722
Qy 1631 CATTTTATACATGAGATCTCTTTGTGTGAATACAGAGATTGACTGCATCCCTTAA 1690
Db 1723 CATTTTATACATGAGATCTCTTTGTGTGAATACAGAGATTGACTGCATCCCTTAA 1781
Qy 1691 AAGAGTTTATGTCCTGACTGTGCTAAATATCTAATTTCCAGATGCTTTGTAGA 1750
Db 1782 AAGAGTTTATGTCCTGACTGTGCTAAATATCTAATTTCCAGATGCTTTGTAGA 1841
Qy 1751 TGACTGAATATTTGTGAGCCACATATTGGAGTTCTAGATTGAGTAATGGCAGAAA 1810
Db 1842 TGACTGAATATTTGTGAGCCACATATTGGAGTTCTAGATTGAGTAATGGCAGAAA 1901
Qy 1811 GGGCATTCTCATTTGAGATGATTAAGTAAGCAACCAACTAGTTCTCGAATTTCTACAGAAA 1870
Db 1902 GGGCATTCTCATTTGAGATGATTAAGTAAGCAACCAACTAGTTCTCGAATTTCTACAGAAA 1961
Qy 1871 GAGGGAATCAGACTGAGAAAGCTGTACATAGACTTGAAGACCAAGACTTTGAATT 1930
Db 1962 GAGGGAATCAGACTGAGAAAGCTGTACATAGACTTGAAGACCAAGACTTTGAATT 2021
Qy 1931 TGGCAGCTGCATGTGTGAGTTATATCACTGCTCTTTCTTCTATGAGTTACAATCTA 1990
Db 2022 TGGCAGCTGCATGTGTGAGTTATATCACTGCTCTTTCTTCTATGAGTTACAATCTA 2081
Qy 1991 TATTTTATGAACTTAAATAAAGAAAATTTACAA 2028
Db 2082 TATTTTATGAACTTAAATAAAGAAAATTTACAA 2119

RESULT 12

US-60-453-050-275
; Sequence 275, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453.050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 275
; LENGTH: 2119
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-050-275

Query Match 30.5%; Score 622; DB 11; Length 2119;
Best Local Similarity 99.4%; Pred. No. 1,4e-298;
Matches 992; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```

OY 1031 GTCAAAAGTTTCATCCAGGGCCACAAAGCTTACAGCCACTCCAAATAAGATGCCATACAT 1090
    |||||||
DB 1123 GTGCAAGATTTCATCCAGGGCCACAAAGCTTACAGCCACTCCAAATAAGATGCCATACAT 1182
OY 1091 GAACTGGAACAGAGAAATTATACCTGTGTGACCTCTGTGATGCAATCATCTATGGG 1150
    |||||||
DB 1183 GAAGCTGAACAAAGAGAAATTATACCTGTGTGACCTCTGTGATGCAATCATCTATGGG 1242
OY 1151 GATCCGGAATGGGCGCCACATAAATCCAAATCCCACTTGAACCACTGAAGAAAGA 1210
    |||||||
DB 1243 GATCCGGAATGGGCGCCACATAAATCCAAATCCCACTTGAACCACTGAAGAAAGA 1302
OY 1211 AGAAGATTGCACTCAGAGCTGTCAACACCATAGAAAGTCAGAGTGTTCGCCAGACTAT 1270
    |||||||
DB 1303 AGAAGATTGCACTCAGAGCTGTCAACACCATAGAAAGTCAGAGTGTTCGCCAGACTAT 1362
OY 1271 AACAAAGAACTTAAAGGAAGGATCCCAAGGGCAGAAATGATCAAGAGCTGAATGCACC 1330
    |||||||
DB 1363 AACAAAGAACTTAAAGGAAGGATCCCAAGGGCAGAAATGATCAAGAGCTGAATGCACC 1422
OY 1331 GTTTAAGAGACTGTCCAGTGGCCCTTTGGAAGGTGTGGGATCCAGTTCAGAGAGGAG 1390
    |||||||
DB 1423 GTTTAAGAGACTGTCCAGTGGCCCTTTGGAAGGTGTGGGATCCAGTTCAGAGAGGAG 1482
OY 1391 GGGTATGTTTGTCTCCAGTCTGGGCAAGAGATGCTATGCGGAATTCCTCATAGCAG 1450
    |||||||
DB 1483 GGGTATGTTTGTCTCCAGTCTGGGCAAGAGATGCTATGCGGAATTCCTCATAGCAG 1542
OY 1451 AAAAGCTCCACACATTTCTTTTGTATGTGTTTAAAGTCTCAGCTTCTATATAGAA 1510
    |||||||
DB 1543 AAAAGCTCCACACATTTCTTTTGTATGTGTTTAAAGTCTCAGCTTCTATATAGAA 1602
OY 1511 ACAGAGGCTTGTGAGCTCCCTGTGTGTGCTATGTCTGTGAATGAGTGAATTCAGGA 1570
    |||||||
DB 1603 ACAGAGGCTTGTGAGCTCCCTGTGTGTGCTATGTCTGTGAATGAGTGAATTCAGGA 1662
OY 1571 AAGCATTTTCTTTCTTTGAACCTTAAAGGTTCTATTTTAAAGCAGACAGATTCGA 1630
    |||||||
DB 1663 AAGCATTTTCTTTCTTTGAACCTTAAAGGTTCTATTTTAAAGCAGACAGATTCGA 1722
OY 1631 CATTTTATACATGAGAGATCTTCTTGTGTGTAATACAGAGATTGATGATCCCTTTAA 1690
    |||||||
DB 1723 CATTTTATACATGAGAGATCTTCTTGTGTGTAATACAGAGATTGATGATCCCTTTAA 1781
OY 1691 AAGAGTTTATGTCCCTGACCTGTGGCTAAATTTATCTAATTTCCAGATGCTTTGTAGA 1750
    |||||||
DB 1782 AAGAGTTTATGTCCCTGACCTGTGGCTAAATTTATCTAATTTCCAGATGCTTTGTAGA 1841
OY 1751 TGACTGAAGTATTTGTGAGCCACATATTGGAGATTCTAGATTGAGTGAATGGCAGAAA 1810
    |||||||
DB 1842 TGACTGAAGTATTTGTGAGCCACATATTGGAGATTCTAGATTGAGTGAATGGCAGAAA 1901
OY 1811 GGGCATCTCCATTTAGATGATTAAGTGAACCAAACTAGTCTCGGAATTTCTACAGAGAA 1870
    |||||||
DB 1902 GGGCATCTCCATTTAGATGATTAAGTGAACCAAACTAGTCTCGGAATTTCTACAGAGAA 1961
OY 1871 GGAGGAATCAGACTGAGGAAGCTGTGACATAGGACTTGAACCAAACTAGTGAAT 1930
    |||||||
DB 1962 GGAGGAATCAGACTGAGGAAGCTGTGACATAGGACTTGAACCAAACTAGTGAAT 2021
OY 1931 TGCGAGCTGCTCAGTGTGAGTATTAATCACTGCTGTCTTTCTATGAGTACAAATCTA 1990
    |||||||
DB 2022 TGCGAGCTGCTCAGTGTGAGTATTAATCACTGCTGTCTTTCTATGAGTACAAATCTA 2081
OY 1991 TATTTTATGAAGTTTAAATTAAGAAAAAATTTTACAA 2028
    |||||||
DB 2082 TATTTTATGAAGTTTAAATTAAGAAAAAATTTTACAA 2119
    |||||||

```

RESULT 13
 US-09-532-315B-3472
 ; Sequence 3472, Application US/09532315B
 ; GENERAL INFORMATION:

```

: APPLICANT: Sellhammer, Jeffrey J.
: APPLICANT: Deleageane, Angelo M.
: APPLICANT: Stuart, Susan G.
: APPLICANT: Stuve, Laura L.
: APPLICANT: Mullahy, Sara J.
: TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES
: FILE REFERENCE: PD-1002 CIP
: CURRENT APPLICATION NUMBER: US/09/532,315B
: PRIOR FILING DATE: 2000-03-24
: PRIOR APPLICATION NUMBER: 07/916,491
: PRIOR FILING DATE: 1992-07-17
: PRIOR APPLICATION NUMBER: 07/977,780
: PRIOR FILING DATE: 1992-11-19
: PRIOR APPLICATION NUMBER: 08/100,523
: PRIOR FILING DATE: 1993-08-03
: PRIOR APPLICATION NUMBER: 09/008,119
: PRIOR FILING DATE: 1998-01-16
: PRIOR APPLICATION NUMBER: 08/196,364
: PRIOR FILING DATE: 1994-02-14
: PRIOR APPLICATION NUMBER: 08/282,991
: PRIOR FILING DATE: 1994-07-28
: PRIOR APPLICATION NUMBER: 08/438,571
: PRIOR FILING DATE: 1995-05-10
: PRIOR APPLICATION NUMBER: 08/179,873
: PRIOR FILING DATE: 1994-01-11
: PRIOR APPLICATION NUMBER: 08/504,732
: PRIOR FILING DATE: 1995-07-20
: PRIOR APPLICATION NUMBER: 08/137,951
: PRIOR FILING DATE: 1993-10-14
: Remaining Prior Application data removed - See file wrapper or PALM.
: NUMBER OF SEQ ID NOS: 42212
: SOFTWARE: PERL Program
: SEQ ID NO: 3472
: LENGTH: 555
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No: hu01286535
US-09-532-315B-3472

Query Match          19.8%; Score 404; DB 6; Length 555;
Best Local Similarity 100.0%; Pred. No. 4,4e-190;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 14
US-09-532-315B-3471
Sequence 3471, Application US/09532315B
GENERAL INFORMATION:
APPLICANT: Sellhamer, Jeffrey J.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullahy, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES
FILE REFERENCE: PD-1002 CIP
CURRENT FILING DATE: 2000-03-24
CURRENT APPLICATION NUMBER: US/09/532,315B
PRIOR APPLICATION NUMBER: 07/916,491
PRIOR FILING DATE: 1992-07-17
PRIOR APPLICATION NUMBER: 07/977,780
PRIOR FILING DATE: 1992-11-19
PRIOR APPLICATION NUMBER: 08/100,523
PRIOR FILING DATE: 1993-08-03
PRIOR APPLICATION NUMBER: 09/008,119
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/196,364
PRIOR FILING DATE: 1994-02-14
PRIOR APPLICATION NUMBER: 08/282,991
PRIOR FILING DATE: 1994-07-28
PRIOR APPLICATION NUMBER: 08/438,571
PRIOR FILING DATE: 1995-05-10
PRIOR APPLICATION NUMBER: 08/179,873
PRIOR FILING DATE: 1994-01-11
PRIOR APPLICATION NUMBER: 08/504,732
PRIOR FILING DATE: 1995-07-20
PRIOR APPLICATION NUMBER: 08/137,951
PRIOR FILING DATE: 1993-10-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 42212
SOFTWARE: PERL Program
SEQ ID NO 3471
LENGTH: 392
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: hu01253362
US-09-532-315B-3471

Query Match 13.2%; Score 269; DB 6; Length 392;
Best Local Similarity 99.7%; Pred. No. 6,7e-123;
Matches 319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1690 AAGAACTTTATGTCCTGACCTGCTGCTAAATTAATCTAATTTCCAGATGCTTTGTAG 1749
DB 69 AAGAACTTTATGTCCTGACCTGCTGCTAAATTAATCTAATTTCCAGATGCTTTGTAG 128
QY 1750 ATGACCTAATTTTGTGACCCACATATTGGGAGTTCTAATTTGATGATGCGAGAA 1809
DB 129 ATGACCTAATTTTGTGACCCACATATTGGGAGTTCTAATTTGATGATGCGAGAA 188
QY 1810 AGGCGCATCTCATTTGAGATGATTAAGTAACCAACTAGTTCGGAATTTTACAGAGA 1869
DB 189 AGGCGCATCTCATTTGAGATGATTAAGTAACCAACTAGTTCGGAATTTTACAGAGA 248
QY 1870 AGAGGGAATCAGACTGAGAGAGCTGTGACATAGGAGTTGAAGCAACCAAGTTTGAAT 1929
DB 249 AGAGGGAATCAGACTGAGAGAGCTGTGACATAGGAGTTGAAGCAACCAAGTTTGAAT 308
QY 1930 TTGCGACTGCTCATGTGTGAGTTATATACATGCTCTTTCTATTGAGTTACAAATCT 1989
DB 309 TTGCGACTGCTCATGTGTGAGTTATATACATGCTCTTTCTATTGAGTTACAAATCT 368
QY 1990 ATATTTTATGAAGTTAA 2009

DB 369 ATATTTTATGAAGTTAA 368
RESULT 15
US-09-532-315B-9521
Sequence 9521, Application US/09532315B
GENERAL INFORMATION:
APPLICANT: Sellhamer, Jeffrey J.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullahy, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES
FILE REFERENCE: PD-1002 CIP
CURRENT FILING DATE: 2000-03-24
CURRENT APPLICATION NUMBER: 07/916,491
PRIOR FILING DATE: 1992-07-17
PRIOR APPLICATION NUMBER: 07/977,780
PRIOR FILING DATE: 1992-11-19
PRIOR APPLICATION NUMBER: 08/100,523
PRIOR FILING DATE: 1993-08-03
PRIOR APPLICATION NUMBER: 09/008,119
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/196,364
PRIOR FILING DATE: 1994-02-14
PRIOR APPLICATION NUMBER: 08/282,991
PRIOR FILING DATE: 1994-07-28
PRIOR APPLICATION NUMBER: 08/438,571
PRIOR FILING DATE: 1995-05-10
PRIOR APPLICATION NUMBER: 08/179,873
PRIOR FILING DATE: 1994-01-11
PRIOR APPLICATION NUMBER: 08/504,732
PRIOR FILING DATE: 1995-07-20
PRIOR APPLICATION NUMBER: 08/137,951
PRIOR FILING DATE: 1993-10-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 42212
SOFTWARE: PERL Program
SEQ ID NO 9521
LENGTH: 293
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: hu00875621
US-09-532-315B-9521

Query Match 13.0%; Score 265; DB 6; Length 293;
Best Local Similarity 100.0%; Pred. No. 6,6e-121;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 GTGGGAGTGCGGCTCAGGGGCTGCAACGAGACCTACCTCTTAGTATTCGCGGCC 103
DB 29 GTGGGAGTGCGGCTCAGGGGCTGCAACGAGACCTACCTCTTAGTATTCGCGGCC 88
QY 104 ACGGGACCGCGCAATTCACAGCTGGCGTTGACAGTGGCGAGCGCTCGCGGTGAGATC 163
DB 89 ACGGGACCGCGCAATTCACAGCTGGCGTTGACAGTGGCGAGCGCTCGCGGTGAGATC 148
QY 164 GTGAGGCTGATCTTCATGACAGGCTCTATGAAGGCTTGAACATCTACCAACCAAGTTTCT 223
DB 149 GTGAGGCTGATCTTCATGACAGGCTCTATGAAGGCTTGAACATCTACCAACCAAGTTTCT 208
QY 224 GCCCAGAGAGAGAGATCTCGCGCACACATGATCAGTTGTGATCCTCTTGAGAC 283
DB 209 GCCCAGAGAGAGAGATCTCGCGCACACATGATCAGTTGTGATCCTCTTGAGAC 268
QY 284 AATTACACAGTGTGAGATTGAGAA 308
DB 269 AATTACACAGTGTGAGATTGAGAA 293

Tue Apr 22 09:11:14 2003

us-09-513-151-3.oli.rnpn

Page 11

Search completed: April 22, 2003, 00:01:33
Job time : 977 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 19:32:57 ; Search time 5186 Seconds

(Without alignments)
9895.067 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041
Sequence: 1 CTCGCATTAACATGCGCTCCG.....TTTACAGAAAAA 2041

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 24791104 seqs, 12571243825 residues

Word size: 0

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

```

1: Pending_Patents_MA_Main:*
2: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US095_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*

```

Result No.	Score	Query Match	Length	ID	Description
1	2041	100.0	2041	US-09-513-151-3	Sequence 3, Appl
2	939	46.0	2108	US-10-133-013-165	Sequence 155, App
3	939	46.0	2129	PCT-US02-07626-152	Sequence 152, App
4	939	46.0	2129	US-10-097-340-152	Sequence 152, App
5	918	45.0	2109	US-60-324-185-33690	Sequence 33690, A
6	823	40.3	2129	PCT-US02-07826-150	Sequence 150, App
7	823	40.3	2129	US-10-097-340-150	Sequence 150, App
8	818	40.1	2101	US-60-112-360-26034	Sequence 26034, A
9	778	38.1	1183	US-09-757-028-874	Sequence 874, App
10	778	38.1	1183	US-10-222-911-874	Sequence 874, App
11	757	37.1	1041	US-09-205-070-8140	Sequence 8140, Ap
12	757	37.1	1041	US-09-340-623-8140	Sequence 8140, Ap
13	757	37.1	1041	US-09-888-888-8140	Sequence 8140, Ap
14	757	37.1	1041	US-09-888-888-8140	Sequence 8140, Ap
15	751	36.8	1749	PCT-US02-18947-1692	Sequence 1692, Ap
16	751	36.8	1749	US-10-172-118-1692	Sequence 1692, Ap
17	697	34.1	1223	US-09-705-256A-6006	Sequence 6006, Ap
18	697	34.1	1223	US-60-164-285-6006	Sequence 6006, Ap
19	696	34.1	1039	US-09-359-822-3465	Sequence 3465, Ap
20	696	34.1	1039	US-09-359-822-3465	Sequence 3465, Ap
21	696	34.1	1039	US-09-919-002-3465	Sequence 3465, Ap

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

22 694 34.0 1858 18 US-09-471-275-2814
 23 638 31.3 1009 17 US-09-399-932-5314
 24 480 23.5 519 25 US-09-652-121-891
 25 473 23.2 717 1 PCT-US01-01239-494
 26 473 23.2 717 1 PCT-US01-01349-166
 27 473 23.2 717 1 PCT-US01-01349-364
 28 473 23.2 717 30 US-09-764-853-166
 29 473 23.2 717 30 US-09-764-853-364
 30 473 23.2 717 30 US-09-764-902-494
 31 473 23.2 717 30 US-09-764-902-166
 32 473 23.2 717 30 US-09-764-902-166
 33 459 22.5 508 25 US-09-652-127-7656
 34 457 22.4 1069 25 US-09-652-124-7628
 35 447 21.9 487 16 US-09-235-076-20244
 36 447 21.9 487 16 US-09-289-768-25904
 37 447 21.9 487 17 US-09-332-782-20244
 38 447 21.9 487 29 US-09-737-223-20244
 39 447 21.9 487 34 US-09-918-995-20244
 40 447 21.9 487 35 US-09-939-397-25904
 41 443 21.7 636 22 US-09-584-852-5059
 42 438 21.5 480 16 US-09-289-768-25903
 43 438 21.5 480 35 US-09-939-397-25903
 44 431 21.1 487 18 US-09-489-036-15715
 45 431 21.1 487 35 US-09-943-143-15715

ALIGNMENTS

Sequence 2814, Ap
 Sequence 5314, Ap
 Sequence 891, App
 Sequence 494, App
 Sequence 166, App
 Sequence 364, App
 Sequence 166, App
 Sequence 494, App
 Sequence 166, App
 Sequence 364, App
 Sequence 7656, App
 Sequence 7628, App
 Sequence 20244, A
 Sequence 25904, A
 Sequence 20244, A
 Sequence 20244, A
 Sequence 20244, A
 Sequence 5059, A
 Sequence 25903, A
 Sequence 25903, A
 Sequence 15715, A
 Sequence 15715, A

RESULT 1
 US-09-513-151-3
 Sequence 3, Application US/09513151
 GENERAL INFORMATION:
 APPLICANT: MCGILL UNIVERSITY
 TITLE OF INVENTION: THE C. ELEGANS gro-1 GENE
 NUMBER OF SEQUENCES: 62
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SMARBY OGILVY RENAULT
 STREET: 1981 McGill College Avenue - Suite 1600
 CITY: Montreal
 STATE: QC
 COUNTRY: Canada
 ZIP: H3A 2Y3
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows
 SOFTWARE: FASTSEQ for Windows Version 2.0b
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/513,151
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/CA98/00803
 FILING DATE: 20-AUG-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: CA 2,210,251
 FILING DATE: 25-AUG-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: CLE, France
 REGISTRATION NUMBER: 4166
 REFERENCE/DOCKET NUMBER: 1770-179"US" FC/9C
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 514 845-7126
 TELEFAX: 514 288-8389
 TELEX:
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2041 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA

US-09-513-151-3
 Query Match 100.0%; Score 2041; DB 19; Length 2041;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCCATTAAGATGGCGCGCTGGCGCTGACAGAGAGTTCCTGTGGGAGTGGGCTGAG 60
 DB 1 CTGCCATTAAGATGGCGCGCTGGCGCTGACAGAGAGTTCCTGTGGGAGTGGGCTGAG 60
 QY 61 GGGCTGCAAGGAGACCTTACCTTGTAGATTCGCGGGCCAGCGGACCGCAAAATC 120
 DB 61 GGGCTGCAAGGAGACCTTACCTTGTAGATTCGCGGGCCAGCGGACCGCAAAATC 120
 QY 121 CACGCTGGCGCTTACCTGAGGAGGCGCTGGCGCTGAGATGTCAGGCGTACATC 180
 DB 121 CACGCTGGCGCTTACCTGAGGAGGCGCTGGCGCTGAGATGTCAGGCGTACATC 180
 QY 181 GCAAGCTTATGAGAGGCTTACATCATCACCAAGAGGTTTCGCCCAAGCAGAGAAAT 240
 DB 181 GCAAGCTTATGAGAGGCTTACATCATCACCAAGAGGTTTCGCCCAAGCAGAGAAAT 240
 QY 241 CTGCCGACACCATATGATGCTTGTGATCCTTGTGACCAATTAACAGATGGTGG 300
 DB 241 CTGCCGACACCATATGATGCTTGTGATCCTTGTGACCAATTAACAGATGGTGG 300
 QY 301 CTTCAGAAATAGACCACTGCTCTGATTAAGATATTTGCCGAGACAAATTCCTAT 360
 DB 301 CTTCAGAAATAGACCACTGCTCTGATTAAGATATTTGCCGAGACAAATTCCTAT 360
 QY 361 TGTGTTGGGAGGAGCAATTTATTAATTTCTGCTGTGAAAGTTCTTCTCAATAC 420
 DB 361 TGTGTTGGGAGGAGCAATTTATTAATTTCTGCTGTGAAAGTTCTTCTCAATAC 420
 QY 421 CAAGCCCCAGAGATGGGAGCTGAGAAAGATTTGACCCGAAAGTGGAGTTGAAAGAG 480
 DB 421 CAAGCCCCAGAGATGGGAGCTGAGAAAGATTTGACCCGAAAGTGGAGTTGAAAGAG 480
 QY 481 GGATGGCTTGTATCTTACAAACGCTTGAAGCGTGGAGCCGAAAGTGGTCCAGCT 540
 DB 481 GGATGGCTTGTATCTTACAAACGCTTGAAGCGTGGAGCCGAAAGTGGTCCAGCT 540
 QY 541 GCATCCACATGACAAACGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
 DB 541 GCATCCACATGACAAACGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
 QY 601 CTCTCATAGTGAATTTCTCATGTCACATACAGAGAGGAGGAGGAGGAGGAGGAGG 660
 DB 601 CTCTCATAGTGAATTTCTCATGTCACATACAGAGAGGAGGAGGAGGAGGAGGAGG 660
 QY 661 TCCTGTGAAGTCTTACACGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 720
 DB 661 TCCTGTGAAGTCTTACACGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 720
 QY 721 TGAGCCCTTGGATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
 DB 721 TGAGCCCTTGGATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
 QY 781 AGATTTTACAGAGGCTTATATATAGAGAGATGTTTGGAAATATGACAGAGATATCA 840
 DB 781 AGATTTTACAGAGGCTTATATATAGAGAGATGTTTGGAAATATGACAGAGATATCA 840
 QY 841 TGGTATCTTCCAAATCAATTTGGCTTCAAGGAATTTCAAGGATCTGATCACTGAGG 900
 DB 841 TGGTATCTTCCAAATCAATTTGGCTTCAAGGAATTTCAAGGATCTGATCACTGAGG 900
 QY 901 ATGCAACACTGAGAGCTAGTACAGCTTCTAAAGAAAGAGGAGGAGGAGGAGGAGG 960
 DB 901 ATGCAACACTGAGAGCTAGTACAGCTTCTAAAGAAAGAGGAGGAGGAGGAGGAGG 960
 QY 961 TGTCTATGCTTAAAGAGGATGATGATGCTGCAAGTGGAGGAGGAGGAGGAGGAGG 1020
 DB 961 TGTCTATGCTTAAAGAGGATGATGATGCTGCAAGTGGAGGAGGAGGAGGAGGAGG 1020


```

OY 1021 TCTTGAATGCTGCAAAAGTTTCATCCAGGGCCACAGCCTACAGCACTCCCAATTAAGAT 1080
    |||||||
Db 1021 TCTTGAATGCTGCAAAAGTTTCATCCAGGGCCACAGCCTACAGCACTCCCAATTAAGAT 1080
OY 1081 GCCATACAAATGAAGCTGAGAAACAAGAGAAAGTTATCACCCTGTGTGACCTCTGTGATCGAAT 1140
    |||||||
Db 1081 GCCATACAAATGAAGCTGAGAAACAAGAGAAAGTTATCACCCTGTGTGACCTCTGTGATCGAAT 1140
OY 1141 CATCATTTGGGATCGCGAATGGGACCGCCACATAAATCCAAATCCCACTTGAAACCAACT 1200
    |||||||
Db 1141 CATCATTTGGGATCGCGAATGGGACCGCCACATAAATCCAAATCCCACTTGAAACCAACT 1200
OY 1201 GAAAGAAAAGAAAGATTGGACCTGAGATGCTGTCAACCCATAGAAATCAGAGTGTTC 1260
    |||||||
Db 1201 GAAAGAAAAGAAAGATTGGACCTGAGATGCTGTCAACCCATAGAAATCAGAGTGTTC 1260
OY 1261 CCCAGACTATTAACAAGAAAGCTTAAAGGAGAGGATCCCAAGGCGAGATGATCAAGACT 1320
    |||||||
Db 1261 CCCAGACTATTAACAAGAAAGCTTAAAGGAGAGGATCCCAAGGCGAGATGATCAAGACT 1320
OY 1321 GAAATGACGCTTTAAGAGACATGTCAGTGCCCTTTGGAAGGTGGTGGGATCCAGTT 1380
    |||||||
Db 1321 GAAATGACGCTTTAAGAGACATGTCAGTGCCCTTTGGAAGGTGGTGGGATCCAGTT 1380
OY 1381 CAGAGGAGAGGGGTATGTTTGTCTCCCGAGTCTGGGCAAGAGTCTATGCGGATTTCTC 1440
    |||||||
Db 1381 CAGAGGAGAGGGGTATGTTTGTCTCCCGAGTCTGGGCAAGAGTCTATGCGGATTTCTC 1440
OY 1441 TGCATAGCAGAAAAGCTCCCAACATTTCTTTGATGTGTTTAAAGTCTCACCGTTCTC 1500
    |||||||
Db 1441 TGCATAGCAGAAAAGCTCCCAACATTTCTTTGATGTGTTTAAAGTCTCACCGTTCTC 1500
OY 1501 TATAATAGAAACAGCAGCTCTTGTGACGCTCTGTGTGGCTGATGTCTGGAATGATG 1560
    |||||||
Db 1501 TATAATAGAAACAGCAGCTCTTGTGACGCTCTGTGTGGCTGATGTCTGGAATGATG 1560
OY 1561 TAGTTCAGGAAAGCAATTTTTTTTCTTTGAACTTAAAGGTTCTATTATTAAAGCAGC 1620
    |||||||
Db 1561 TAGTTCAGGAAAGCAATTTTTTTTCTTTGAACTTAAAGGTTCTATTATTAAAGCAGC 1620
OY 1621 ACAGATTCACATTTTATACATGAGATCTTCTTGTGCTGAATACAGAGATTGACTGC 1680
    |||||||
Db 1621 ACAGATTCACATTTTATACATGAGATCTTCTTGTGCTGAATACAGAGATTGACTGC 1680
OY 1681 ATCCCTTTAAAGAGATTTTATGTCCTGACTCTGGCTAAATATTCTAATTTCCAGATG 1740
    |||||||
Db 1681 ATCCCTTTAAAGAGATTTTATGTCCTGACTCTGGCTAAATATTCTAATTTCCAGATG 1740
OY 1741 CTTTGTGATGATGAGTATTTGTGAGCCCATATTGGGAGTCTGATTTGAGTGA 1800
    |||||||
Db 1741 CTTTGTGATGATGAGTATTTGTGAGCCCATATTGGGAGTCTGATTTGAGTGA 1800
OY 1801 TGGCAGGAAAGGGCCATCTCCATGAGATGATTAAAGTACCAAACTACTTCGGAAT 1860
    |||||||
Db 1801 TGGCAGGAAAGGGCCATCTCCATGAGATGATTAAAGTACCAAACTACTTCGGAAT 1860
OY 1861 CTACAGAGAGAGGAGAACTCAGACTGAGAGAAAGCTGTGACATGAGCTTGAGACCAAGA 1920
    |||||||
Db 1861 CTACAGAGAGAGGAGAACTCAGACTGAGAGAAAGCTGTGACATGAGCTTGAGACCAAGA 1920
OY 1921 CTTTAAATTTGCGAGCTGCTCATGTGTGAGTATTATACAGTCTCTTTCTATTGAGT 1980
    |||||||
Db 1921 CTTTAAATTTGCGAGCTGCTCATGTGTGAGTATTATACAGTCTCTTTCTATTGAGT 1980
OY 1981 TACAATCTATATTTTATGAAAGTTTAAATTAAGAAAAATTTTACAGAAAAA 2040
    |||||||
Db 1981 TACAATCTATATTTTATGAAAGTTTAAATTAAGAAAAATTTTACAGAAAAA 2040
OY 2041 A 2041
    |||||||
Db 2041 A 2041

```

```

RESULT 2
US-10-133-013-165
: Sequence 165, Application US/10133013
: GENERAL INFORMATION:
: APPLICANT: Astromoff, Anna
: APPLICANT: Bandman, Olga
: APPLICANT: Cocks, Benjamin G.
: TITLE OF INVENTION: GENES ASSOCIATED WITH VASCULAR DISEASE
: FILE REFERENCE: PA-0049 US
: CURRENT APPLICATION NUMBER: US/10/133,013
: PRIOR FILING DATE: 2002-04-25
: PRIOR APPLICATION NUMBER: 60/287,067
: NUMBER OF SEQ ID NOS: 271
: SOFTWARE: PERL Program
: SEQ ID NO 165
: LENGTH: 2108
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc-feature
: OTHER INFORMATION: Incyte ID No: 977391.15
US-10-133-013-165

Query Match          46.0%; Score 939; DB 40; Length 2108;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 939; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGCCATTAAGATGCGCTCCGTCGCGGCTGACAGACAGTCTCTGTGGGCACTGGCTCAG 60
Db 1 CTGCCATTAAGATGCGCTCCGTCGCGGCTGACAGACAGTCTCTGTGGGCACTGGCTCAG 60
OY 61 GGGCGTGAAGGAGGACCTTACCTCTGTAGTATTCGCGGGCCAGCGGACCGGCAATC 120
Db 61 GGGCGTGAAGGAGGACCTTACCTCTGTAGTATTCGCGGGCCAGCGGACCGGCAATC 120
OY 121 CAGCGTGGCGTTGACAGTACAGGCGGCTGCGGCGTGAAGTGTACGCGTACTCCAT 180
Db 121 CAGCGTGGCGTTGACAGTACAGGCGGCTGCGGCGTGAAGTGTACGCGTACTCCAT 180
OY 181 GCAGGTCTATGAAGGCTTACAGATCATCACCAAGAGTTTGTGCCCAAGACAGAGAT 240
Db 181 GCAGGTCTATGAAGGCTTACAGATCATCACCAAGAGTTTGTGCCCAAGACAGAGAT 240
OY 241 CTGCCGCGACCATGATGATGCTTGTGGATCTCTGTGACCAATTTACAGTGTGGA 300
Db 241 CTGCCGCGACCATGATGATGCTTGTGGATCTCTGTGACCAATTTACAGTGTGGA 300
OY 301 CTTTCAAAATAGAGCACTGCTGTGATTGAAGATATTTTGTGCCGAGATCAAAATTCAT 360
Db 301 CTTTCAAAATAGAGCACTGCTGTGATTGAAGATATTTTGTGCCGAGATCAAAATTCAT 360
OY 361 TGTGTGGAGAGAACCAATTTATACATTTGAATCTGTCTGTGGAAGTTCTTGTCAATAC 420
Db 361 TGTGTGGAGAGAACCAATTTATACATTTGAATCTGTCTGTGGAAGTTCTTGTCAATAC 420
OY 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTATTTACACCAAAAGTGGAGCTTGAAGA 480
Db 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTATTTACACCAAAAGTGGAGCTTGAAGA 480
OY 481 GGATGGTCTTACTTCAACAAAGCTTAAGCCAGGTGAGCCAGAAATGGCTGCCAAGCT 540
Db 481 GGATGGTCTTACTTCAACAAAGCTTAAGCCAGGTGAGCCAGAAATGGCTGCCAAGCT 540
OY 541 GCATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGGAT 600
Db 541 GCATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGGAT 600
OY 601 CTCTCATAGTGAATTTCTCATGCTACATACGGAAGAGTGTGTGCTCCCTTGGAGG 660
Db 601 CTCTCATAGTGAATTTCTCATGCTACATACGGAAGAGTGTGTGCTCCCTTGGAGG 660

```

OY 661 TCCTCTGAAGTTCCTTAACCCCTTGCTGCTTGGCTTCATGCTGACAGGAGTTCCTAGA 720
DB 661 TCCTCTGAAGTTCCTTAACCCCTTGCTGCTTGGCTTCATGCTGACAGGAGTTCCTAGA 720
OY 721 TGAGCGCTTGATTAAGAGGGTGATGATGATGCTTGGCTGCTTGGAGGAACTAG 780
DB 721 TGAGCGCTTGATTAAGAGGGTGATGATGATGCTTGGCTGCTTGGAGGAACTAG 780
OY 781 AGATTTTCACAGCGCTTATTCAGAGAAATGTTTGGAAATAGCCAGGACTATCAACA 840
DB 781 AGATTTTCACAGCGCTTATTCAGAGAAATGTTTGGAAATAGCCAGGACTATCAACA 840
OY 841 TGGTATCTTCCATCAATTTGGCTTCAAGGATTTACAGAGTACCTGATCAGTGAAGGAAA 900
DB 841 TGGTATCTTCCATCAATTTGGCTTCAAGGATTTACAGAGTACCTGATCAGTGAAGGAAA 900
OY 901 ATGCACACTGGAGACTAGTAACGACGCTTCTTAAGAAAG 939
DB 901 ATGCACACTGGAGACTAGTAACGACGCTTCTTAAGAAAG 939

RESULT 3
PCT-US02-07826-152
Sequence 152, Application PC/TUS0207826
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc. et al.
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins for The Identification,
Title Of Invention: Assessment, Prevention, and Therapy of Ovarian Cancer
FILE REFERENCE: MRI-030PC
CURRENT APPLICATION NUMBER: PCT/US02/07826
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 152
LENGTH: 2129
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US02-07826-152

Query Match 46.0%; Score 939; DB 1; Length 2129;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 939; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTGCCATAAATAGCGCTCCCTGGCGGCTGCACAGCAGCTTCTGTGGCGAGTGGCTCAG 60
DB 5 CTGCCATAAATAGCGCTCCCTGGCGGCTGCACAGCAGCTTCTGTGGCGAGTGGCTCAG 64
OY 61 GGGCCTGCAAGGACCTTACTCTTGTAGTATCTCGGGGGCGACGGGCAACGCAAAATC 120
DB 65 GGGCCTGCAAGGACCTTACTCTTGTAGTATCTCGGGGGCGACGGGCAACGCAAAATC 124
OY 121 CACGCTGGCGCTTGACGTAGCGGCGGCTCGGCGGTGAGATGTCAGGCGTACTTCAT 180
DB 125 CACGCTGGCGCTTGACGTAGCGGCGGCGGCTCGGCGGTGAGATGTCAGGCGTACTTCAT 184
OY 181 GCGAGGCTATGAGGCGCTAGACATCATCACAAAGGTTTCTGCCCAAGAGCAGAGAT 240
DB 185 GCGAGGCTATGAGGCGCTAGACATCATCACAAAGGTTTCTGCCCAAGAGCAGAGAT 244

OY 241 CTGCCGAGACACATGATACCTTTGGTATCTCTTGTGTGACCAATTAACAGTGTGGA 300
DB 245 CTGCCGAGACACATGATACCTTTGGTATCTCTTGTGTGACCAATTAACAGTGTGGA 304
OY 301 CTTGAGAAATAGCAACTGCTGATTTGAAGATATATTTCCCGACAGCAAAATTTCTAT 360
DB 305 CTTGAGAAATAGCAACTGCTGATTTGAAGATATATTTCCCGACAGCAAAATTTCTAT 364
OY 361 TGGTGGGAGGAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420
DB 365 TGGTGGGAGGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 424
OY 421 CAACCCCGAGAGATGGGCACTGAGAAAGTATGACCGAAAGTGAAGCTTGAAGA 480
DB 425 CAACCCCGAGAGATGGGCACTGAGAAAGTATGACCGAAAGTGAAGCTTGAAGA 484
OY 481 GGATGCTCTTGTACTTACCAAAACGCTTAAGCAGGTGAGCCCAAAATGGCTGCAAGCT 540
DB 485 GGATGCTCTTGTACTTACCAAAACGCTTAAGCAGGTGAGCCCAAAATGGCTGCAAGCT 544
OY 541 GCATCCACATGACAAACGCAAGTGCCAGAGCTTCAAGTTTGAAGAAACAGGAAT 600
DB 545 GCATCCACATGACAAACGCAAGTGCCAGAGCTTCAAGTTTGAAGAAACAGGAAT 604
OY 601 CTCTCATAGTGAATTTCTCATGCTCAACATAGCGAAAGGTGCTGCCCTTGAGG 660
DB 605 CTCTCATAGTGAATTTCTCATGCTCAACATAGCGAAAGGTGCTGCCCTTGAGG 664
OY 661 TCCTCTGAAGTTCCTTAACCCCTTGCTGCTTGGCTTCATGCTGACAGGAGTTCCTAGA 720
DB 665 TCCTCTGAAGTTCCTTAACCCCTTGCTGCTTGGCTTCATGCTGACAGGAGTTCCTAGA 724
OY 721 TGAGCGCTTGATTAAGAGGGTGATGATGATGCTTGGCTGCTTGGAGGAACTAG 780
DB 725 TGAGCGCTTGATTAAGAGGGTGATGATGATGCTTGGCTGCTTGGAGGAACTAG 784
OY 781 AGATTTTCACAGCGCTTATTCAGAGAAATGTTTGGAAATAGCCAGGACTATCAACA 840
DB 785 AGATTTTCACAGCGCTTATTCAGAGAAATGTTTGGAAATAGCCAGGACTATCAACA 844
OY 841 TGGTATCTTCCATCAATTTGGCTTCAAGGATTTACAGAGTACCTGATCAGTGAAGGAAA 900
DB 845 TGGTATCTTCCATCAATTTGGCTTCAAGGATTTACAGAGTACCTGATCAGTGAAGGAAA 904
OY 901 ATGCACACTGGAGACTAGTAACGACGCTTCTTAAGAAAG 939
DB 905 ATGCACACTGGAGACTAGTAACGACGCTTCTTAAGAAAG 943

RESULT 4
US-10-097-340-152
Sequence 152, Application US/10097340
GENERAL INFORMATION:
APPLICANT: John MONAHAN
APPLICANT: Manjula GANNANARAPU
APPLICANT: Sebastian HOESCH
APPLICANT: Shubhangi KAMATKAR
APPLICANT: Steve G. KOVATS
APPLICANT: Rachel E. MEYERS
APPLICANT: Michael MORRISSEY
APPLICANT: Peter OLANDT
APPLICANT: Ami SEN
APPLICANT: Peter VEIBY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. BAST, Jr.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHMANDT
APPLICANT: Xumei ZHAO
APPLICANT: Karen GLATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins for The Identification,
Title Of Invention: Assessment, Prevention, and Therapy of Ovarian Cancer
FILE REFERENCE: MRI-030
CURRENT APPLICATION NUMBER: US/10/097,340

Query Match	40.3%	Score 823:	DB 1:	Length 2129:
Best Local Similarity	100.0%:	Pred. No. 0:		
Matches 823:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
QY	117	AATCAGCGTGGCGTTGCACCTAGGCGAGCGGCGTGGGCGTGAGATGTCACGCGTACT	176	
Db	121	AATCAGCGTGGCGTTGCACCTAGGCGAGCGGCGTGGGCGTGAGATGTCACGCGTACT	180	
QY	177	CCATGCAGGTCCTATGAAGGCGTTAGACATCATCCAAACAAGTTTCTGCCAAGACAGA	236	
Db	181	CCATGCAGGTCCTATGAAGGCGTTAGACATCATCCAAACAAGTTTCTGCCAAGACAGA	240	
QY	237	GAATTCGCGCGGACCCCATGATGACGTTTGTGGATCTCTTGTGACCAATTACAGTGG	296	
Db	241	GAATTCGCGCGGACCCCATGATGACGTTTGTGGATCTCTTGTGACCAATTACAGTGG	300	
QY	297	TGGACTTCAGAAATAGAGCAACTGCTCTGATTTGAAGATATATTGCGCCAGCAAAATTC	356	
Db	301	TGGACTTCAGAAATAGAGCAACTGCTCTGATTTGAAGATATATTGCGCCAGCAAAATTC	360	
QY	357	CTATTGTGTGGGAGGACCAATTATTTACATTTGAATCTCTCTGGAAGTTCTTGCA	416	
Db	361	CTATTGTGTGGGAGGACCAATTATTTACATTTGAATCTCTCTGGAAGTTCTTGCA	420	
QY	417	ATACCAAGCGCCCGAGAGATGGGACATGACGAAGTATGACCCGAAATGGAGCTTGAA	476	
Db	421	ATACCAAGCGCCCGAGAGATGGGACATGACGAAGTATGACCCGAAATGGAGCTTGAA	480	
QY	477	AGGAGGATGGTCTTGTACTTTCACAAACGCGTTAGCCAGGTGGACCCAGAAATGGCTGCA	536	
Db	481	AGGAGGATGGTCTTGTACTTTCACAAACGCGTTAGCCAGGTGGACCCAGAAATGGCTGCA	540	
QY	537	AGCTGATCCACATGACAAAGCAAGTGGCCAGGAGCTTGCAACTTTTGAAGAAACAG	596	
Db	541	AGCTGATCCACATGACAAAGCAAGTGGCCAGGAGCTTGCAACTTTTGAAGAAACAG	600	
QY	597	GAATCTCTCATGTAGTAATTTCTCATCGCAACATAGGAAAGGTGGTGGTCCCTTG	656	
Db	601	GAATCTCTCATGTAGTAATTTCTCATCGCAACATAGGAAAGGTGGTGGTCCCTTG	660	
QY	657	GAGGTCCTCTGAGTTCTCTAACCCCTTGCACTTTGGGCTTCATGCTACGACGACAGTTC	716	
Db	661	GAGGTCCTCTGAGTTCTCTAACCCCTTGCACTTTGGGCTTCATGCTACGACGACAGTTC	720	
QY	717	TGATGAGCGGCTTGATTAAGAGGGTGAATGACATGCTTGGCTGGGCGCTTTGGAGAAC	776	
Db	721	TGATGAGCGGCTTGATTAAGAGGGTGAATGACATGCTTGGCTGGGCGCTTTGGAGAAC	780	
QY	777	TAAAGATTTTTCACAGACGCTATATATCAGAGATGTTTGGGAAATATTCAGAGCATATC	836	
Db	781	TAAAGATTTTTCACAGACGCTATATATCAGAGATGTTTGGGAAATATTCAGAGCATATC	840	
QY	837	AACATGTAATCTTCCCAATCAATTGGCTTCAAGGAATTTTCAGAGTACTGATCACTAGAG	896	
Db	841	AACATGTAATCTTCCCAATCAATTGGCTTCAAGGAATTTTCAGAGTACTGATCACTAGAG	900	
QY	897	GAAATGCAACCTGGAGACTAGTACCAAGCTTCTAAAGAAAG	939	
Db	901	GAAATGCAACCTGGAGACTAGTACCAAGCTTCTAAAGAAAG	943	
RESULT 7				
US-10-097-340-150				
: Sequence 150, Application US/10097340				
: GENERAL INFORMATION:				
: APPLICANT: John MONAHAN				
: APPLICANT: Manjula GANNAVAPU				
: APPLICANT: Sebastian HOERSCH				
: APPLICANT: Shubhangi KAMATKAR				
: APPLICANT: Steve G. KOVATS				
: APPLICANT: Rachel E. MEYERS				
: APPLICANT: Michael MORRISSEY				

```
APPLICANT: Peter OLANDT
APPLICANT: Ami SEN
APPLICANT: Peter VEIBY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. BAST, Jr.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHMANDT
APPLICANT: Xumei ZHAO
APPLICANT: Karen GLATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
FILE REFERENCE: MRI-030
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: US/10/097,340
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 150
LENGTH: 2129
TYPE: DNA
ORGANISM: Homo sapiens
US-10-097-340-150

Query Match          40.3%; Score 823; DB 39; Length 2129;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 823; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 AATCCAGCGTGGCGCTTGCAGCTAGGCGCGGCTGCGCGGTGAGATGCTGACGCCCTGACT 176
DB 121 AATCCAGCGTGGCGCTTGCAGCTAGGCGCGGCTGCGCGGTGAGATGCTGACGCCCTGACT 180
OY 177 CCATGCGAGGTCTATGAGAGGCTTACATCATCAGCAACAGGTTTCCCAAGAGAGA 236
DB 181 CCATGCGAGGTCTATGAGAGGCTTACATCATCAGCAACAGGTTTCCCAAGAGAGA 240
OY 237 GAATCTGCGGCGACCATGATCAGCTTTGTGATCCTCTTGTGACCAATTACACAGTGG 296
DB 241 GAATCTGCGGCGACCATGATCAGCTTTGTGATCCTCTTGTGACCAATTACACAGTGG 300
OY 297 TGCAGCTTCAGAAATAGAGCAACTGCTGTGATTAAGATATATTTGCCGAGACAAATTC 356
DB 301 TGCAGCTTCAGAAATAGAGCAACTGCTGTGATTAAGATATATTTGCCGAGACAAATTC 360
OY 357 CTATTTGTTGGGAGAGCAACCAATTATATCATTTGTAATCTCTCTGGAAGTTCTTGCA 416
DB 361 CTATTTGTTGGGAGAGCAACCAATTATATCATTTGTAATCTCTCTGGAAGTTCTTGCA 420
OY 417 ATACCAAGCCCCAGAGATGGGCACTGAGAAAGTATGACCAAGAAAGTGGAGCTTGAAA 476
DB 421 ATACCAAGCCCCAGAGATGGGCACTGAGAAAGTATGACCAAGAAAGTGGAGCTTGAAA 480
OY 477 AGAGAGATGCTTGTACTTACAAACGCTTAAGCCAGGTGAGACCCAGCAAAATGCTGCCA 536
DB 481 AGAGAGATGCTTGTACTTACAAACGCTTAAGCCAGGTGAGACCCAGCAAAATGCTGCCA 540
OY 537 AGCTGATTCACATGACAAAGCGCAAGTGGCCAGAGCTTCCAAAGTTTGAAGAAAGAG 596
DB 541 AGCTGATTCACATGACAAAGCGCAAGTGGCCAGAGCTTCCAAAGTTTGAAGAAAGAG 600
OY 597 GAATCTCTCATAGTAATTTCTCCATCTCTCAACATATGAGAAAGTGGTGTGCCCTTG 656
```

```
DB 601 GAATCTCTCATAGTAATTTCTCCATCTCTCAACATATGAGAAAGTGGTGTGCCCTTG 660
OY 657 GAGTCTCTGAGAGTTCTCTTAACCTTGATCTTGGCTTCAATGCTGACAGCAGATTTC 715
DB 661 GAGTCTCTGAGAGTTCTCTTAACCTTGATCTTGGCTTCAATGCTGACAGCAGATTTC 720
OY 717 TAGATGAGCGCTTGATTAAGAGGTTGATGATGATGCTTGTGCTGGCTTGGAGAAC 776
DB 721 TAGATGAGCGCTTGATTAAGAGGTTGATGATGATGCTTGTGCTGGCTTGGAGAAC 780
OY 777 TAAGAGATTTTCACACAGCCTTAATATACAGAGAAATTTGGGAAATAGCCAGACTATC 836
DB 781 TAAGAGATTTTCACACAGCCTTAATATACAGAGAAATTTGGGAAATAGCCAGACTATC 840
OY 837 AACATGATATCTCCATCAATGAGCTTCAAGGAATTTCAAGATACCTGATCAGTACAG 896
DB 841 AACATGATATCTCCATCAATGAGCTTCAAGGAATTTTCACAGATACCTGATCAGTACAG 900
OY 897 GAAATGACACCTGAGAGCTAGTAACCAAGCTTCTAAGAAAG 939
DB 901 GAAATGACACCTGAGAGCTAGTAACCAAGCTTCTAAGAAAG 943

RESULT 8
US-60-172-360-26034
Sequence 26034, Application US/60172360
GENERAL INFORMATION:
APPLICANT: Morris, Macdonald
APPLICANT: Lai, Preclt
APPLICANT: Diep, Dinh
TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms
TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
FILE REFERENCE: GX-0007 P
CURRENT APPLICATION NUMBER: US/60/172,360
CURRENT FILING DATE: 1999-12-16
NUMBER OF SEQ ID NOS: 29838
SOFTWARE: PERL Program
SEQ ID NO 26034
LENGTH: 2101
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 330917.13
NAME/KEY: unsure
LOCATION: 2098
OTHER INFORMATION: a, t, c, g, or other
US-60-172-360-26034

Query Match          40.1%; Score 818; DB 61; Length 2101;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1088; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

OY 939 GACCTGTCCTCATTTGTCCTCCCTGCTATGCTTAGAGATATGATGTCTGAGTGGG 998
DB 1007 GACCTGTCCTCATTTGTCCTCCCTGCTATGCTTAGAGATATGATGTCTGAGTGGG 1066
OY 999 AGAGCTCTGTTCTTGACCTGCTTGAATTCGTGCAAGTTTATCCAGGCCCAAGC 1058
DB 1067 AAGAGTCTGTTCTTGACCTGCTTGAATTCGTGCAAGTTTATCCAGGCCCAAGC 1126
OY 1059 CTACAGCAGCTCCATTAAGATGCGATCATGAGAGCTGAGAACAGAGAGATTATCACC 1118
DB 1127 CTACAGCAGCTCCATTAAGATGCGATCATGAGAGCTGAGAACAGAGAGATTATCACC 1186
OY 1119 TGTGTGACCTCTGTGATGAAATCATATTTGGGATGCGGAATGGCGACGACATAAAT 1178
DB 1187 TGTGTGACCTCTGTGATGAAATCATATTTGGGATGCGGAATGGCGACGACATAAAT 1246
OY 1179 CCAATCCCACTTGACCACTGAAGAAAGAGAGTTGGAGCTCAGATCTGTCAACA 1238
```

Db 1247 CCAATCCCACTTGAAACCACTGAGAGAAAAGAGATTGAGCTGATGCTGTCAACA 1306
Oy 1239 CCATAGAAAGTCAGAGTGTTCCTCCAGACTATATACAAAGAACTAAAGGAAGGATCCC 1298
Db 1307 CCATAGAAAGTCAGAGTGTTCCTCCAGACTATATACAAAGAACTAAAGGAAGGATCCC 1366
Oy 1299 CAGGCGAATGATGATCAAGAGCTGAAATGCAAGCTTTAGAGACATGCTCCAGTGGCTTTG 1358
Db 1367 CAGGCGAATGATGATCAAGAGCTGAAATGCAAGCTTTAGAGACATGCTCCAGTGGCTTTG 1426
Oy 1359 GAAAGGTGGGGGAGTCCAGTTCCAGAGGAGGGGTATGTTGTCTCCAGTCTGGGCAA 1418
Db 1427 GAAAGGTGGGGGAGTCCAGTTCCAGAGGAGGGGTATGTTGTCTCCAGTCTGGGCAA 1486
Oy 1419 AGAGAGTCTATGCGGAATTCCTGATACAGAGAAAGCTCCCACTTTTCTTTTATGAT 1478
Db 1487 AGAGAGTCTATGCGGAATTCCTGATACAGAGAAAGCTCCCACTTTTCTTTTATGAT 1546
Oy 1479 GGTTTTAAAGCTCACGTTCTCTATATATAGAAAACAGCAGGTCTTGTACCTCTGTGTG 1538
Db 1547 GGTTTTAAAGCTCACGTTCTCTATATATAGAAAACAGCAGGTCTTGTACCTCTGTGTG 1606
Oy 1539 GGTGATGTGCTGGAATGATGATGCTGAGAAAGCATTTTCTTTTGAACCTTAA 1598
Db 1607 GGTGATGTGCTGGAATGATGATGCTGAGAAAGCATTTTCTTTTGAACCTTAA 1666
Oy 1599 AGGTCTCTATTTTAAAGCAGACAGATCCACATTTTATACATGAGATCTCTCTTGT 1658
Db 1667 AGGTCTCTATTTTAAAGCAGACAGATCCACATTTTATACATGAGATCTCTCTTGT 1726
Oy 1659 GGTGAATACAGAGATTGATGCTGATCCCTTTAAAGAGTTTATGTCCCTGACTGGCT 1718
Db 1727 GGTGAATACAGAGATTGATGCTGATCCCTTTAAAGAGTTTATGTCCCTGACTGGCT 1785
Oy 1719 AAAATTATCTAATTTCCAGATGCTTTTGTAGTGAAGTATTTGTGAGCCACATATT 1778
Db 1786 AAAATTATCTAATTTCCAGATGCTTTTGTAGTGAAGTATTTGTGAGCCACATATT 1845
Oy 1779 GGGAGTCTAGATTGATGATGATGCGAGAAAGGCGCATTCCTGATGATGATTAAGTG 1838
Db 1846 GGGAGTCTAGATTGATGATGATGCGAGAAAGGCGCATTCCTGATGATGATTAAGTG 1905
Oy 1839 AACCAACTAGTCTCGGAATTTCTACAGAGAAGGAGGAATCAGACTAGAGAACTGTGA 1898
Db 1906 AACCAACTAGTCTCGGAATTTCTACAGAGAAGGAGGAATCAGACTAGAGAACTGTGA 1965
Oy 1899 CATGGAGTCTGATGATGATGATGCGAGAAAGGCGCATTCCTGATGATGATTAAGTG 1958
Db 1966 CATGGAGTCTGATGATGATGATGCGAGAAAGGCGCATTCCTGATGATGATTAAGTG 2025
Oy 1959 CACTGCTGCTCTTCTATGATGATGATGCGAGAAAGGCGCATTCCTGATGATGATTAAGTG 2018
Db 2026 CACTGCTGCTCTTCTATGATGATGATGCGAGAAAGGCGCATTCCTGATGATGATTAAGTG 2085
Oy 2019 AAATTTACAGA 2030
Db 2086 AAATTTACAGA 2097

RESULT 9
US-09-757-028-874
Sequence 874, Application US/09757028
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PM001
CURRENT APPLICATION NUMBER: US/09/757,028
CURRENT FILING DATE: 2001-01-09
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
NUMBER OF SEQ ID NOS: 2660

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 874
LENGTH: 1183
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1172)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1177)
OTHER INFORMATION: n equals a,t,g, or c
US-09-757-028-874

Query Match 38.1%; Score 778; DB 29; Length 1183;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1098; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Oy 939 GACCTGTGTCCTTGAACCTGCTGTAATGTCGAAGTTTCATCCAGGGCCACAGC 1058
Db 35 GACCTGTGTCCTTGAACCTGCTGTAATGTCGAAGTTTCATCCAGGGCCACAGC 154
Oy 999 AGAGTCTGTTCTTGAACCTGCTGTAATGTCGAAGTTTCATCCAGGGCCACAGC 1058
Db 95 AGAGTCTGTTCTTGAACCTGCTGTAATGTCGAAGTTTCATCCAGGGCCACAGC 154
Oy 1059 CTACAGCCACTCCCAATTAAGATGCTATCAATGAGCTGAGAACAGAGATTATCAC 1118
Db 155 CTACAGCCACTCCCAATTAAGATGCTATCAATGAGCTGAGAACAGAGATTATCAC 214
Oy 1119 TGTGTGACCTCTGTATGATGATCATATGCGGATGCGGAATGGCGACGACATMAAT 1178
Db 215 TGTGTGACCTCTGTATGATGATCATATGCGGATGCGGAATGGCGACGACATMAAT 274
Oy 1179 CCAATCCCACTGAGAACCAAGCAAGAAAGAAAGAGTTGAGCTCAGATCTGTCAACA 1238
Db 275 CCAATCCCACTGAGAACCAAGCAAGAAAGAAAGAGTTGAGCTCAGATCTGTCAACA 334
Oy 1239 CCATAGAAAGTCAGAGTGTTCCTCCAGACTATTAACAAGAACTTAAGAGAAAGGATCCC 1298
Db 335 CCATAGAAAGTCAGAGTGTTCCTCCAGACTATTAACAAGAACTTAAGAGAAAGGATCCC 394
Oy 1299 CAGGCGAATGATGATGATGAGCTGAAATGCAAGCTGTTTAAAGACATGTCAGTGGCTTTG 1358
Db 395 CAGGCGAATGATGATGATGAGCTGAAATGCAAGCTGTTTAAAGACATGTCAGTGGCTTTG 454
Oy 1359 GAAAGGTGGGGGAGTCCAGTTCCAGAGGAGGGGTATGTTGTCTCCAGTGGGCAA 1418
Db 455 GAAAGGTGGGGGAGTCCAGTTCCAGAGGAGGGGTATGTTGTCTCCAGTGGGCAA 514
Oy 1419 AGAGTCTATGCGGAATTTCTGATAGCAAGAAAGCTCCACATTTTCTTTTGAATGT 1478
Db 515 AGAGTCTATGCGGAATTTCTGATAGCAAGAAAGCTCCACATTTTCTTTTGAATGT 574
Oy 1479 GGTTTTAAAGTCTCACGTTCTCTATATAGAAACAGCAGGTCTTGTACGCTTGTGTG 1538
Db 575 GGTTTTAAAGTCTCACGTTCTCTATATAGAAACAGCAGGTCTTGTACGCTTGTGTG 634
Oy 1539 GGTGATGTGCTGGAATGATGATGATGCGAGAAAGGCGCATTCCTGATGATGATTAAGTG 1598
Db 635 GGTGATGTGCTGGAATGATGATGATGCGAGAAAGGCGCATTCCTGATGATGATTAAGTG 694
Oy 1599 AGGTCTATATTTAAAGCAGACAGATTCACATTTTATACATGAGGATCTCTTGTGT 1658
Db 695 AGGTCTATATTTAAAGCAGACAGATTCACATTTTATACATGAGGATCTCTTGTGT 754
Oy 1659 GGTGAATACAGAGATTGATGATGATGCGGTTTAAAGAGAGTTTATGCTCCAGTCTGGCT 1718
Db 755 GGTGAATACAGAGATTGATGATGATGCGGTTTAAAGAGAGTTTATGCTCCAGTCTGGCT 813
Oy 1719 AAAATTATCTAATTTCCAGATGCTTTTGTAGAGTCAAGTATTTGTGAGCCACATATT 1778
Db 814 AAAATTATCTAATTTCCAGATGCTTTTGTAGAGTCAAGTATTTGTGAGCCACATATT 873

QY 1601 GTCTATTATTAAGCAGACAGATCCCATTTTATACATGAGGATCTCTTTGCG 1660
|||||
Db 441 GTCTATTATTAAAGAGCAGACAGATTCACATTTTATACATGAGGATCTCTTTGCG 382
QY 1661 TGAATACAGAGATGAGTGCATCCCTTTAAAGAAATTTATGCTCCCTGACTGCTAA 1720
|||||
Db 381 TGAATACAGAGATGAGTGCATCCCTTTAAAGAAATTTATGCTCCCTGACTGCTAA 322
QY 1721 AATTATCTAATTTCCCATGCTTTTGTAGATGACTGAAGTATTTGTAGCCACATTTGG 1780
|||||
Db 321 AATTATCTAATTTCCCATGCTTTTGTAGATGACTGAAGTATTTGTAGCCACATTTGG 262
QY 1781 GAGTTCTAGATTTGAGTGAATGGCAGAAAGGCCATCTCCATGAGATGATTAAGTAA 1840
|||||
Db 261 GAGTTCTAGATTTGAGTGAATGGCAGAAAGGCCATCTCCATGAGATGATTAAGTAA 202
QY 1841 CCAAACTAGTTCTCGAATTTCTACAGAAAGGAGGAATCGACTGAGGAAGCTGTGACA 1900
|||||
Db 201 CCAAACTAGTTCTCGAATTTCTACAGAAAGGAGGAATCGACTGAGGAAGCTGTGACA 142
QY 1901 TAGACTTTGAAGCAAGACTTTGAAATTTGCGAGCTGCTCATGTTGATTTATCA 1960
|||||
Db 141 TAGACTTTGAAGCAAGACTTTGAAATTTGCGAGCTGCTCATGTTGATTTATCA 82
QY 1961 CTGCTGCTTTCTATTGATGATCAAAATCTATTTTATTGAGTTTAAATAAGAAAAA 2020
|||||
Db 81 CTGCTGCTTTCTATTGATGATCAAAATCTATTTTATTGAGTTTAAATAAGAAAAA 22
QY 2021 ATTACCAAGAAAAA 2041
|||||
Db 21 ATTACCAAGAAAAA 1
RESULT 13
US-09-888-8140/c
Sequence 8140, Application US/0989888
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
FILE REFERENCE: 20411-748CON1
CURRENT APPLICATION NUMBER: US/09/898, 888
CURRENT FILING DATE: 2001-07-03
PRIOR APPLICATION NUMBER: 09/340, 623
PRIOR FILING DATE: 1999-06-28
NUMBER OF SEQ ID NOS: 45207
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8140
LENGTH: 1041
TYPE: DNA
ORGANISM: Homo sapiens
US-09-888-8140
Query Match 37.1% Score 757; DB 33; Length 1041;
Best Local Similarity 99.6% Pred. No. 0;
Matches 1037; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 1243 AGAAGTCAGAGTGTTCCTCCAGACTATACAAAGAACCTAAAGGAGGATCCCCAG 1302
|||||
Db 801 AGAAGTCAGAGTGTTCCTCCAGACTATACAAAGAACCTAAAGGAGGATCCCCAG 742
QY 1303 GCAGAAATGATCAAGAGCTGAATGACAGCTTTAAAGACATGTCAGTGCCTTTGGAA 1362
|||||
Db 741 GCAGAAATGATCAAGAGCTGAATGACAGCTTTAAAGACATGTCAGTGCCTTTGGAA 682
QY 1363 GGTGCTGGGATCCAGTTCAGAGAGGAGGGGATGTTGTCCTCCAGTCCGGGCAAGAA 1422
|||||
Db 681 GGTGCTGGGATCCAGTTCAGAGAGGAGGGGATGTTGTCCTCCAGTCCGGGCAAGAA 622
QY 1423 GTGCTATGCGGAATTTCTCTCATAGCAGAAAGCTCCACCATTTTCTTTGATGTGTT 1482
|||||
Db 621 GTGCTATGCGGAATTTCTCTCATAGCAGAAAGCTCCACCATTTTCTTTGATGTGTT 562
QY 1483 TTAAGTCTCAGCTTCTATTAATAGAAACGACAGTCTGTCAGTCTGTCGCTG 1542
|||||
Db 561 TTAAGTCTCAGCTTCTATTAATAGAAACGACAGTCTGTCAGTCTGTCGCTG 502
QY 1543 ATGCTGCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1600
|||||
Db 501 ATGCTGCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 442
QY 1601 GTCTATTATTAAGCAGACAGATTCACATTTTATACATGAGGATCTCTTTGCG 1660
|||||
Db 441 GTCTATTATTAAGCAGACAGATTCACATTTTATACATGAGGATCTCTTTGCG 382
QY 1661 TGAATACAGAGATGAGTGCATCCCTTTAAAGAAATTTATGCTCCCTGACTGCTAA 1720
|||||
Db 381 TGAATACAGAGATGAGTGCATCCCTTTAAAGAAATTTATGCTCCCTGACTGCTAA 322
QY 1721 AATTATCTAATTTCCCATGCTTTTGTAGATGACTGAAGTATTTGTAGCCACATTTGG 1780
|||||
Db 321 AATTATCTAATTTCCCATGCTTTTGTAGATGACTGAAGTATTTGTAGCCACATTTGG 262
QY 1781 GAGTTCTAGATTTGAGTGAATGGCAGAAAGGCCATCTCCATGAGATGATTAAGTAA 1840
|||||
Db 261 GAGTTCTAGATTTGAGTGAATGGCAGAAAGGCCATCTCCATGAGATGATTAAGTAA 202
QY 1841 CCAAACTAGTTCTCGAATTTCTACAGAAAGGAGGAATCGACTGAGGAAGCTGTGACA 1900
|||||
Db 201 CCAAACTAGTTCTCGAATTTCTACAGAAAGGAGGAATCGACTGAGGAAGCTGTGACA 142
QY 1901 TAGACTTTGAAGCAAGACTTTGAAATTTGCGAGCTGCTCATGTTGATTTATCA 1960
|||||
Db 141 TAGACTTTGAAGCAAGACTTTGAAATTTGCGAGCTGCTCATGTTGATTTATCA 82
QY 1961 CTGCTGCTTTCTATTGATGATCAAAATCTATTTTATTGAGTTTAAATAAGAAAAA 2020
|||||
Db 81 CTGCTGCTTTCTATTGATGATCAAAATCTATTTTATTGAGTTTAAATAAGAAAAA 22
QY 2021 ATTACCAAGAAAAA 2041
|||||
Db 21 ATTACCAAGAAAAA 1
RESULT 14
US-09-888-8140/c
Sequence 8140, Application US/0989888
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
FILE REFERENCE: 20411-748CON1
CURRENT APPLICATION NUMBER: US/09/898, 888
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/340, 623
PRIOR FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: US 09/205, 070
PRIOR FILING DATE: 1998-12-03
NUMBER OF SEQ ID NOS: 45207

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 19:34:21 ; Search time 85 Seconds

(Without alignments)
7363.850 Million cell updates/sec

Title: US-09-513-151-3
Perfect score: 2041

Sequence: 1 CTCGCATAGATGCGCTCCG.....TTTACAGAAAAA 2041

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size: 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	23	1.1 87350	3 US-08-781-891-79	Sequence 79, Appl
2	23	1.1 87543	4 US-09-791-211-3	Sequence 3, Appl
3	23	1.0 98844	4 US-09-791-211-10	Sequence 10, Appl
4	19	0.9 1314	4 US-09-134-001C-2505	Sequence 2505, Ap
5	19	0.9 1659	4 US-09-152-060-24	Sequence 24, Appl
6	19	0.9 49136	4 US-09-422-869-1	Sequence 1, Appl
7	18	0.9 293	1 US-08-222-177A-49	Sequence 49, Appl
8	18	0.9 618	4 US-09-289-349-2	Sequence 2, Appl
9	18	0.9 657	4 US-09-134-001C-2195	Sequence 2195, Ap
10	18	0.9 1319	2 US-08-484-933B-17	Sequence 17, Appl
11	18	0.9 1319	2 US-08-484-158B-17	Sequence 17, Appl
12	18	0.9 1319	2 US-08-484-596A-17	Sequence 17, Appl
13	18	0.9 1319	2 US-08-480-150A-17	Sequence 17, Appl
14	18	0.9 1319	3 US-08-458-731-17	Sequence 17, Appl
15	18	0.9 1319	3 US-08-149-223A-17	Sequence 17, Appl
16	18	0.9 1761	3 US-09-043-830-2	Sequence 2, Appl
17	18	0.9 2100	1 US-08-485-718-10	Sequence 10, Appl
18	18	0.9 2100	2 US-08-484-530-56	Sequence 56, Appl
19	18	0.9 2100	2 US-08-827-618A-56	Sequence 56, Appl
20	18	0.9 2100	3 US-08-483-952A-56	Sequence 56, Appl
21	18	0.9 2100	4 US-08-476-501-56	Sequence 56, Appl
22	18	0.9 2100	6 5475086-5	Patent No. 5475086
23	18	0.9 2249	2 US-08-494-624-1	Sequence 1, Appl
24	18	0.9 2370	1 US-08-117-907-1	Sequence 1, Appl
25	18	0.9 2370	3 US-08-453-040-1	Sequence 1, Appl
26	18	0.9 2756	1 US-08-187-793-1	Sequence 1, Appl
27	18	0.9 3318	1 US-08-187-793-3	Sequence 3, Appl

C 28	18	0.9	5973	4	US-09-245-041-4	Sequence 4, Appl
C 29	18	0.9	6623	2	US-08-687-080-68	Sequence 68, Appl
C 30	18	0.9	21234	4	US-09-810-671-3	Sequence 3, Appl
C 31	18	0.9	32042	4	US-09-245-281-44	Sequence 44, Appl
C 32	18	0.9	168575	4	US-09-426-290-17	Sequence 1, Appl
C 33	18	0.9	176373	3	US-09-128-155-17	Sequence 17, Appl
C 34	18	0.8	40	4	US-09-306-290-31	Sequence 31, Appl
C 35	17	0.8	260	1	US-08-594-031-45	Sequence 45, Appl
C 36	17	0.8	260	1	US-08-594-031-137	Sequence 137, Appl
C 37	17	0.8	260	1	US-08-594-031-144	Sequence 144, Appl
C 38	17	0.8	272	2	US-08-454-557C-100	Sequence 100, Appl
C 39	17	0.8	272	2	US-08-340-426D-100	Sequence 100, Appl
C 40	17	0.8	272	2	US-08-450-673C-100	Sequence 100, Appl
C 41	17	0.8	272	5	PCT-US95-17111A-100	Sequence 100, Appl
C 42	17	0.8	324	4	US-08-651-155B-22	Sequence 22, Appl
C 43	17	0.8	452	3	US-09-033-055A-2	Sequence 2, Appl
C 44	17	0.8	496	4	US-08-978-289-3	Sequence 3, Appl
C 45	17	0.8	604	4	US-09-370-838-175	Sequence 175, Appl

ALIGNMENTS

RESULT 1

US-08-781-891-79
Sequence 79, Application US/08781891

Patent No. 6090620

GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui

APPLICANT: Yu, Chang-En

APPLICANT: Oshima, Junko

APPLICANT: Mulligan, John T.

APPLICANT: Schellenberg, Gerald D.

TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO

NUMBER OF SEQUENCES: 209

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,891

FILING DATE: 27-DEC-1996

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: No. 6090620tenburg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 240052.419

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 79:

SEQUENCE CHARACTERISTICS:

LENGTH: 87350 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-781-891-79

Query Match 1.1%: Score 23; DB 3; Length 87350;

Best Local Similarity 100.0%; Pred. No. 0.18;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2019 AATTTACAGAAAAA 2041
|||||

Db 51760 AAATTACAGAAAAA 51782

RESULT 2
US-09-791-211-3

Sequence 3, Application US/09791211

Patent No. 6448080

GENERAL INFORMATION:

APPLICANT: Donna T. Ward

APPLICANT: Andrew T. Walt

TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION

FILE REFERENCE: RTS-0205

CURRENT APPLICATION NUMBER: US/09/791,211

CURRENT FILING DATE: 2001-02-23

NUMBER OF SEQ ID NOS: 90

SEQ ID NO 3

LENGTH: 87543

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: unsure

LOCATION: 7421

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 7427

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 11609

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 12605

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 12742

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 29370

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 29422

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 29979

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 29980

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 30136

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 30140

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 31206

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 31592

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 33095

OTHER INFORMATION: unknown

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 34072

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 36816

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 39020

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 42164

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 42459

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 46808

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 46826

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 47291

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 52786

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 52787

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 53384

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 54684

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 59215

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 59235

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 59242

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 63290

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 66614

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 68660

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 68697

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 68718

OTHER INFORMATION: unknown

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 68733

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 68739

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure
LOCATION: 79134
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 79198
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 86336
OTHER INFORMATION: unknown
OTHER INFORMATION: unknown
US-09-791-211-3

Query Match 1.18; Score 23; DB 4; Length 87543;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2019 AAATTACAGAAAAAAA 2041
|||||
DB 51953 AAATTACAGAAAAAAA 51975

RESULT 3

US-09-791-211-10/c
Sequence 10, Application US/09791211
Patent No. 6448080
GENERAL INFORMATION:
APPLICANT: Donna T. Ward
TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
FILE REFERENCE: RTS-0205
CURRENT APPLICATION NUMBER: US/09/791,211
CURRENT FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 90
SEQ ID NO 10
LENGTH: 98844
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: 24962
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 64383
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 65468
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 65469
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 65470
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 65471
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 87130
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 89049
OTHER INFORMATION: unknown
OTHER INFORMATION: unknown
US-09-791-211-10

Query Match

Best Local Similarity 1.0%; Score 21; DB 4; Length 98844;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2019 AAATTACAGAAAAAAA 2039
|||||
DB 36142 AAATTACAGAAAAAAA 36122

RESULT 4

US-09-134-001C-2505
Sequence 2505, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 2505
LENGTH: 1314
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2505

Query Match 0.9%; Score 19; DB 4; Length 1314;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1595 TTAAGGTTCTATTATTA 1613
|||||
DB 1274 TTAAGGTTCTATTATTA 1292

RESULT 5

US-09-152-060-24/c
Sequence 24, Application US/09152060
Patent No. 6448230
GENERAL INFORMATION:
APPLICANT: Rosen et al
TITLE OF INVENTION: 28 Human Secreted Proteins
FILE REFERENCE: P2003p1.US
CURRENT APPLICATION NUMBER: US/09/152,060
CURRENT FILING DATE: 1998-09-11
EARLIER APPLICATION NUMBER: PCT/US98/04858
EARLIER FILING DATE: 1998-03-12
EARLIER APPLICATION NUMBER: 60/040,762
EARLIER FILING DATE: 1997-03-14
EARLIER APPLICATION NUMBER: 60/040,710
EARLIER FILING DATE: 1997-03-14
EARLIER APPLICATION NUMBER: 60/050,934
EARLIER FILING DATE: 1997-05-30
EARLIER APPLICATION NUMBER: 60/048,100
EARLIER FILING DATE: 1997-05-30
EARLIER APPLICATION NUMBER: 60/048,357
EARLIER FILING DATE: 1997-05-30
EARLIER APPLICATION NUMBER: 60/048,189
EARLIER FILING DATE: 1997-05-30
EARLIER APPLICATION NUMBER: 60/057,765
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/068,368
EARLIER FILING DATE: 1997-12-19
NUMBER OF SEQ ID NOS: 118
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 24
LENGTH: 1659
TYPE: DNA
ORGANISM: Homo sapiens
US-09-152-060-24

Query Match

Best Local Similarity 0.9%; Score 19; DB 4; Length 1659;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 388 TGAATCTGCTTGAGAA 406
|||||
Db 311 TGAATCTGCTTGAGAA 293

RESULT 6

US-09-422-869-1/C
Sequence 1, Application US/09422869
Patent No. 6235481
GENERAL INFORMATION:
APPLICANT: POLONSKY, KENNETH S.
APPLICANT: HORIKAWA, YUKIO
APPLICANT: ODA, NAOHISA
APPLICANT: COX, NANCY J.
APPLICANT: SREENAN, SEAMUS
APPLICANT: ZHOU, YUN-PING
APPLICANT: OTANI, KENICHI
APPLICANT: HANIS, CRAIG L.
APPLICANT: BELL, GRAEME I.
TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
FILE REFERENCE: ARCD:307
CURRENT APPLICATION NUMBER: US/09/422,869
CURRENT FILING DATE: 1999-10-21
EARLIER APPLICATION NUMBER: 60/134,175
EARLIER FILING DATE: 1999-05-13
NUMBER OF SEQ. ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 49136
TYPE: DNA
ORGANISM: Human
US-09-422-869-1

Query Match 0.9%; Score 19; DB 4; Length 49136;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1091 GAAGCTGAGAACAGAGAA 1109
|||||
Db 35847 GAAGCTGAGAACAGAGAA 35829

RESULT 7

US-08-222-177A-49/C
Sequence 49, Application US/08222177A
Patent No. 5382979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dc-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelstor Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.

REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 293 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

INDIVIDUAL ISOLATE: Caucasian

TISSUE TYPE: Blood

IMMEDIATE SOURCE:

CLONE: Mcd122

POSITION IN GENOME:

CHROMOSOME/SEGMENT: 5q

FEATURE:

NAME/KEY: repeat_region

LOCATION: 41..74

OTHER INFORMATION: /rpt_type="tandem"

OTHER INFORMATION: /rpt_family="(dc-da)n.(dc-dt)n"

OTHER INFORMATION: /citation="(12)"

FEATURE:

NAME/KEY: misc_feature

LOCATION: 4..23

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence=EXPERIMENTAL

OTHER INFORMATION: /standard_name="PCR primer"

OTHER INFORMATION: /citation="(11)"

FEATURE:

NAME/KEY: misc_feature

LOCATION: complement (204..223)

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence=EXPERIMENTAL

OTHER INFORMATION: /standard_name="PCR primer"

OTHER INFORMATION: /citation="(11)"

FEATURE:

NAME/KEY: misc_feature

LOCATION: 1..293

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence=EXPERIMENTAL

OTHER INFORMATION: /standard_name="Only one strand sequenced"

PUBLICATION INFORMATION:

AUTHORS: Weber, J. L.

AUTHORS: Polymeropoulos, M. H.

AUTHORS: May, P. E.

AUTHORS: Kwitek, A. E.

AUTHORS: Xiao, H.

AUTHORS: McPherson, J. D.

AUTHORS: Masmuth, J. J.

TITLE: Mapping of human chromosome 5 microsatellite

TITLE: polymorphisms

JOURNAL: Genomics

DATE: 1991

PUBLICATION INFORMATION:

AUTHORS: Weber, James L.

AUTHORS: May, Paula E.

TITLE: Abundant Class of Human DNA Polymorphisms

TITLE: Which Can Be Typed Using the Polymerase Chain

TITLE: Reaction

JOURNAL: Am. J. Hum. Genet.

VOLUME: 44

PAGES: 388-396

DATE: 1989

US-08-222-177A-49

Query Match 0.9%; Score 18; DB 1; Length 293;
Best Local Similarity 100.0%; Pred. No. 46;

Matches 18: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2019 AATTTCACAGAAAAA 2036
|||||
DB 197 AATTTCACAGAAAAA 180

RESULT 8

US-09-289-349-2/C
Sequence 2, Application US/09289349
Patent No. 6277574
GENERAL INFORMATION:
APPLICANT: Walker, Michael, G.
APPLICANT: Volkmuth, Wayne
APPLICANT: Klinger, Tod, M.
APPLICANT: Azimzal, Yalda
APPLICANT: Yue, Henry
TITLE OF INVENTION: GENES ASSOCIATED WITH DISEASES OF THE KIDNEY
FILE REFERENCE: PB-0010 US
CURRENT APPLICATION NUMBER: US/09/289,349
CURRENT FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PERL Program
SEQ ID NO 2
LENGTH: 618
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: 3534377C71
US-09-289-349-2

Query Match 0.9%; Score 18; DB 4; Length 618;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 18: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2024 TACAGAAAAA 2041
|||||
DB 578 TACAGAAAAA 561

RESULT 9

US-09-134-001C-2195
Sequence 2195, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 2195
LENGTH: 657
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2195

Query Match 0.9%; Score 18; DB 4; Length 657;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 18: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2008 AATAAGAAAAATTTA 2025
|||||
DB 210 AATAAGAAAAATTTA 227

RESULT 10
US-08-484-993B-17

Sequence 17, Application US/08484993B
Patent No. 5837497

GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immuncontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,993B
FILING DATE: 09-NOV-1993

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341

FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.

REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653

TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:

LENGTH: 1319 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO

ANTI-SENSE: NO
ORIGINAL SOURCE:

ORGANISM: Felis domesticus
DEVELOPMENTAL STAGE: Juvenile

HAPLOTYPE: Diploidy
TISSUE TYPE: Ovary

CELL TYPE: oocyte
FEATURE:

NAME/KEY: CDS
LOCATION: 26..1297

US-08-484-993B-17

Query Match 0.9%; Score 18; DB 2; Length 1319;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 18: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1702 TGTCCCTGACTGGCTA 1719
|||||
DB 1206 TGTCCCTGACTGGCTA 1223

RESULT 11
US-08-484-158B-17

Sequence 17, Application US/08484158B
Patent No. 5976545
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.

APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Pharmaceutical Compositions for
TITLE OF INVENTION: Immunoccontraception
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,158B
FILING DATE: 07-JUNE-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 09-NOV-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-92
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 32794
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ. ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1319 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Felis domesticus
DEVELOPMENTAL STAGE: Juvenile
HAPLOTYPE: Diploidy
TISSUE TYPE: Ovary
CELL TYPE: Oocyte
FEATURE:
NAME/KEY: CDS
LOCATION: 26..1297
US-08-484-158B-17

Query Match 0.9%; Score 18; DB 2; Length 1319;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1702 TGTCCCTGACTCTGGCTA 1719
|||||
Db 1206 TGTCCCTGACTCTGGCTA 1223

RESULT 12
US-08-484-596A-17
Sequence 17, Application US/08484596A
Patent No. 5981228
GENERAL INFORMATION:

APPLICANT: Hsu, Kuang T., Jeffrey D.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunoccontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,596A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 11-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ. ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1319 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Felis domesticus
DEVELOPMENTAL STAGE: Juvenile
HAPLOTYPE: Diploidy
TISSUE TYPE: Ovary
CELL TYPE: Oocyte
FEATURE:
NAME/KEY: CDS
LOCATION: 26..1297
US-08-484-596A-17

Query Match 0.9%; Score 18; DB 2; Length 1319;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1702 TGTCCCTGACTCTGGCTA 1719
|||||
Db 1206 TGTCCCTGACTCTGGCTA 1223

RESULT 13
US-08-480-150A-17
Sequence 17, Application US/08480150A
Patent No. 5989550
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunoccontraception

NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,150A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/149,223
FILING DATE: 09-NOV-1993
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1319 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Felis domesticus
DEVELOPMENTAL STAGE: Juvenile
HAPLOTYPE: Diploidy
TISSUE TYPE: Ovary
CELL TYPE: Oocyte
FEATURE:
NAME/KEY: CDS
LOCATION: 26..1297
US-08-480-150A-17
Query Match 0.9%; Score 18; DB 2; Length 1319;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1702 TGTCCTGACTGTGCTA 1719
|||||
Db 1206 TGTCCTGACTGTGCTA 1223
RESULT 14
US-08-458-731-17
Sequence 17, Application US/08458731
Patent No. 6001599
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,731
FILING DATE: 09-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1319 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Felis domesticus
DEVELOPMENTAL STAGE: Juvenile
HAPLOTYPE: Diploidy
TISSUE TYPE: Ovary
CELL TYPE: Oocyte
FEATURE:
NAME/KEY: CDS
LOCATION: 26..1297
US-08-458-731-17
Query Match 0.9%; Score 18; DB 3; Length 1319;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1702 TGTCCTGACTGTGCTA 1719
|||||
Db 1206 TGTCCTGACTGTGCTA 1223
RESULT 15
US-08-149-223A-17
Sequence 17, Application US/08149223A
Patent No. 6027727
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois

COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/149,223A
FILING DATE: 09-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1319 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Felis domesticus
DEVELOPMENTAL STAGE: Juvenile
HAPLOTYPE: Diploidy
TISSUE TYPE: Ovary
CELL TYPE: Oocyte
FEATURE:
NAME/KEY: CDS
LOCATION: 26..1297
US-08-149-223A-17

Query Match 0.98; Score 18; DB 3; Length 1319;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 1702 TGTCCCTGACTCTGGCTA 1719
Db 1206 TGTCCCTGACTCTGGCTA 1223

Search completed: April 21, 2003, 23:50:13
Job time : 393 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 19:21:06 ; Search time 480 Seconds
(without alignments)
9575.685 million cell updates/sec

Title: US-09-513-151-3
Perfect score: 2041
Sequence: 1 CTCGCATAGATGCGCTCCG.....TTTACAGAAAAA 2041

Scoring table: OLIGO-NUC
Gap 60.0, Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

N.Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2041	100.0	2041	20	AAK36073 Human homologue of
2	896	43.9	2130	24	ABN59708 Novel human coding
3	473	23.2	717	22	ABA06500 Human CDNA SEQ ID
4	473	23.2	717	22	ABA06598 Human CDNA SEQ ID
5	473	23.2	717	22	AAK41268 cDNA encoding nove
6	241	11.8	355	24	ABL68710 Kidney cancer rela
7	175	8.6	300	21	AAA00911 Human colon cancer
8	145	7.1	457	22	ABA42839 Human breast cell
9	145	7.1	457	22	ABA53262 Human foetal liver

C 10	145	7.1	457	22	AAK01519	Human brain expres
C 11	145	7.1	457	22	AAK26966	Human bone marrow
C 12	145	7.1	457	22	AAI11574	Probe #1507 for ge
C 13	145	7.1	457	22	AAI32862	Probe #1548 used t
C 14	138	6.8	138	22	ABA47966	Human breast cell
C 15	138	6.8	138	22	ABA55854	Human foetal liver
C 16	138	6.8	138	22	AAK14265	Human brain expres
C 17	138	6.8	138	22	AAK39988	Human bone marrow
C 18	138	6.8	138	22	AAI20787	Probe #10720 for g
C 19	138	6.8	138	22	AAI46014	Probe #14700 used
C 20	60	2.9	60	24	ABN34729	Human spliced tran
C 21	59	2.9	582	23	ABV49413	Human prostate exp
C 22	40	2.0	545	23	ABV19646	Human prostate exp
C 23	23	1.1	2001	23	AAK88985	DNA encoding novel
C 24	23	1.1	4557	23	AAK79149	DNA encoding novel
C 25	23	1.1	6881	22	AAI92597	Human polynucleoti
C 26	23	1.1	7313	23	AAK81337	DNA encoding novel
C 27	23	1.1	7488	22	AAK87243	Human immune/haema
C 28	23	1.1	7488	23	AAK87244	Human immune/haema
C 29	23	1.1	11087	23	AAK87437	DNA encoding novel
C 30	23	1.1	11378	23	AAK87336	DNA encoding novel
C 31	23	1.1	13234	23	AAK82685	DNA encoding novel
C 32	23	1.1	15879	23	AAK84109	DNA encoding novel
C 33	23	1.1	33780	22	AAH24652	Nucleotide sequenc
C 34	22	1.1	87350	18	AAK83003	Human WPN genomic
C 35	22	1.1	190	21	AAK26737	Human secreted pro
C 36	22	1.1	885	22	AAK78227	Human immune/haema
C 37	22	1.1	1118	22	AAI36412	Human musculoskele
C 38	22	1.1	1398	22	AAI36415	Human immune/haema
C 39	22	1.1	1398	22	AAK66112	Human immune/haema
C 40	22	1.1	6225	23	ABL10424	Drosophila melanog
C 41	22	1.1	6225	23	ABL06640	Drosophila melanog
C 42	22	1.1	6399	23	ABL17248	Drosophila melanog
C 43	22	1.1	19802	23	ABL12950	Drosophila melanog
C 44	22	1.1	33376	22	AAK66113	Human immune/haema
C 45	22	1.1	33357	24	ABN85746	Human transporter

ALIGNMENTS

RESULT 1
ID AAK36073
AAK36073 standard: DNA; 2041 BP.

AC AAK36073:
15-JUL-1999 (first entry)

DE Human homologue of gro-1, referred to as hgro-1.

XX gro-1 operon: gro-1 gene; gop-1 gene; gop-2 gene; gop-2 gene;
KW hap-1 gene; cancer; aging; longevity; tumour formation;
KW physiological clock; ss.

XX Homo sapiens.
OS WO9910482-A1.
PM 04-MAR-1999.
PD 20-AUG-1998; 98MO-CA00803.
PF 25-AUG-1997; 97CA-2210251.
PR (UYMC-) UNIV MCGILL.
PA Barnes T, Hekimi S, Lakowski B, Lemieux J;
PI WPI: 1999-190615/16.
DR Molecular identity of the gro-1 gene - useful for cancer diagnosis
PT and/or prognosis, and where compounds affecting encoded proteins are

PT useful for enhancing longevity of a host and inhibiting tumour
PT formation
XX
PS Claim 8; Fig 8; 93bp; English.

XX
CC The present sequence represents the human homologue of gro-1, and
CC is referred to as hgro-1. The specification describes the five genes
CC of the Caenorhabditis elegans gro-1 operon (AXX3607). The operon
CC contains the gro-1 gene (AXX36072), the gop-1 gene (AXX36074), the gop-2
CC gene (AXX36075), the gop-2 gene (AXX36075), and the hap-1 gene
CC (AXX36077). The gro-1 gene can be used in a method for the diagnosis
CC and/or prognosis of cancer in a patient. Transgenic mice containing a
CC gene knock-out of a murine gene homologue of the gro-1 gene are useful
CC as models of aging and cancer. The proteins encoded by the genes are
CC useful for identifying compounds that affect the enzymatic activity
CC of these proteins. In order to enhance longevity of a host and inhibit
CC tumour formation, the gro-1 gene, together with the gop-1, gop-2,
CC gop-3 and hap-1 genes enables study of a physiological clock.

XX
SQ Sequence 2041 BP: 589 A; 421 C; 502 G; 529 T; 0 other;

Query Match 100.0%; Score 2041; DB 20; Length 2041;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCCAATAAGATGGCGCGCGGCGCTGCACAGCAGTTCCTGTGGGAGTGGGCTCAG 60
DB 1 CTGCCAATAAGATGGCGCGCGGCGCTGCACAGCAGTTCCTGTGGGAGTGGGCTCAG 60
QY 61 GGGCTTCGCAAGCGACCTTACCTCTTGTAGTATCTCGGGCCACGGGCAATC 120
DB 61 GGGCTTCGCAAGCGACCTTACCTCTTGTAGTATCTCGGGCCACGGGCAATC 120
QY 121 CAGGCTGGGCTTGCAGTACGCGGCGCTGGGCGGTGAGATGTCAGCCCTGATCCAT 180
DB 121 CAGGCTGGGCTTGCAGTACGCGGCGCTGGGCGGTGAGATGTCAGCCCTGATCCAT 180
QY 181 GCAGTCTATGAAGGCTTACATCATCACCAACAAGCTTTCGCCAAGAGCAGAGAT 240
DB 181 GCAGTCTATGAAGGCTTACATCATCACCAACAAGCTTTCGCCAAGAGCAGAGAT 240
QY 241 CTGCGGCGACACATGATGAGCTTGTGATCTCTTGTGACCAATTCACAGTGTGA 300
DB 241 CTGCGGCGACACATGATGAGCTTGTGATCTCTTGTGACCAATTCACAGTGTGA 300
QY 301 CTTCAGAAATAGCAACATGCTCTGATGTAAGATATATTTGCCGAGACAAATTTCTAT 360
DB 301 CTTCAGAAATAGCAACATGCTCTGATGTAAGATATATTTGCCGAGACAAATTTCTAT 360
QY 361 TGTGTGGGAGGAACCAATTAATTAATCTGCTGTGAAAGTTCTTGTCAATAC 420
DB 361 TGTGTGGGAGGAACCAATTAATTAATCTGCTGTGAAAGTTCTTGTCAATAC 420
QY 421 CAAGCCCGAGAGATGGGCACTGAGAAATGATTTGACCGAAAGTGGAGCTTGAAGAAG 480
DB 421 CAAGCCCGAGAGATGGGCACTGAGAAATGATTTGACCGAAAGTGGAGCTTGAAGAAG 480
QY 481 GATGCTCTTGTACTTCAAAAGCGCTAAGCCAGGTGAGCCGGAATGGCTCCCAAGCT 540
DB 481 GATGCTCTTGTACTTCAAAAGCGCTAAGCCAGGTGAGCCGGAATGGCTCCCAAGCT 540
QY 541 GCATCCACATGACAAACGCAAGTGGCGAGAGCTTGAAGTTTGTGAAGAAACAGAGAT 600
DB 541 GCATCCACATGACAAACGCAAGTGGCGAGAGCTTGAAGTTTGTGAAGAAACAGAGAT 600
QY 601 CTCTCATAGTGAATTTCTCATGTCACATACGAGAAAGTGGTGTCTTCTTGGAGG 660
DB 601 CTCTCATAGTGAATTTCTCATGTCACATACGAGAAAGTGGTGTCTTCTTGGAGG 660
QY 661 TCTCTGAAGTCTCTAACCCCTTGCATCTTGGCTTCATGCTGACACGAGCTTCTAGA 720
DB 661 TCTCTGAAGTCTCTAACCCCTTGCATCTTGGCTTCATGCTGACACGAGCTTCTAGA 720

QY 721 TGAGCGCTTGATTAAGGGGTGATGATCATGCTTGTGCTGGGCTCTTGGAGCACTAAC 780
DB 721 TGAGCGCTTGATTAAGGGGTGATGATCATGCTTGTGCTGGGCTCTTGGAGCACTAAC 780
QY 781 AGATTTTACAGACGCTATATATCAGAGAAATGTTTGGAAATATAGCCAGGACTATACA 840
DB 781 AGATTTTACAGACGCTATATATCAGAGAAATGTTTGGAAATATAGCCAGGACTATACA 840
QY 841 TGGTATCTTCAATCAATTTGGCTTCAAGAAATTTACAGATACCTGATCACTAGGAGAA 900
DB 841 TGGTATCTTCAATCAATTTGGCTTCAAGAAATTTACAGATACCTGATCACTAGGAGAA 900
QY 901 ATGCACCTGAGACCTAGTATACCAAGCTTCTTAAGAAAGACCTGTGCTTGTCCCCC 960
DB 901 ATGCACCTGAGACCTAGTATACCAAGCTTCTTAAGAAAGACCTGTGCTTGTCCCCC 960
QY 961 TGTCTATGCTTGAAGTATCTGATGCTGCAAGTGGAGAGATCTTGTGTAACCTGC 1020
DB 961 TGTCTATGCTTGAAGTATCTGATGCTGCAAGTGGAGAGATCTTGTGTAACCTGC 1020
QY 1021 TCTTGAATGTCGCAAGTTCATCCAGGGCCACAAAGCTTACAGCCATCCAAATAAGAT 1080
DB 1021 TCTTGAATGTCGCAAGTTCATCCAGGGCCACAAAGCTTACAGCCATCCAAATAAGAT 1080
QY 1081 GCCATACATGAGCTGAGAACAGAGAAATATATCATCTGTGTACTCTGTATGAT 1140
DB 1081 GCCATACATGAGCTGAGAACAGAGAAATATATCATCTGTGTACTCTGTATGAT 1140
QY 1141 CATCATTTGGGATGCGCAATAGGCGAGGCGCATATAAATCCAAATCCACTTGAACCACT 1200
DB 1141 CATCATTTGGGATGCGCAATAGGCGAGGCGCATATAAATCCAAATCCACTTGAACCACT 1200
QY 1201 GAAGAAAGAAAGATTTGAGTCTGATGCTGCAACACCATAGAAAGTACAGTGTTC 1260
DB 1201 GAAGAAAGAAAGATTTGAGTCTGATGCTGCAACACCATAGAAAGTACAGTGTTC 1260
QY 1261 CCCAGACTATTAACAAAGAACTTAAAGGAGATCCCGGCGAGAAATGATTAAGAGCT 1320
DB 1261 CCCAGACTATTAACAAAGAACTTAAAGGAGATCCCGGCGAGAAATGATTAAGAGCT 1320
QY 1321 GAAATGCAAGCTTTAAGAGACATGTCAGTGGCTTTGGAAGAGTGGTGGGATTCAGTT 1380
DB 1321 GAAATGCAAGCTTTAAGAGACATGTCAGTGGCTTTGGAAGAGTGGTGGGATTCAGTT 1380
QY 1381 CAGAGAGGAGGGTATGTTGCTCCCACTGTGGCAAGAGAGTCTATGCGAATTC 1440
DB 1381 CAGAGAGGAGGGTATGTTGCTCCCACTGTGGCAAGAGAGTCTATGCGAATTC 1440
QY 1441 TGCATAGCAGAAAGCTCCACCATTTCTTTGATGTGTTTAAAGTCTCACGTTCTC 1500
DB 1441 TGCATAGCAGAAAGCTCCACCATTTCTTTGATGTGTTTAAAGTCTCACGTTCTC 1500
QY 1501 TATATATAGAAACAGCAGGCTTGTGAGCTCTTGTGCTGATGCTGTGAAATGATG 1560
DB 1501 TATATATAGAAACAGCAGGCTTGTGAGCTCTTGTGCTGATGCTGTGAAATGATG 1560
QY 1561 TAGTTAGAGAAAGCATTTTCTTTTCTTGAACCTTAAAGTCTATTAATAAGCAGC 1620
DB 1561 TAGTTAGAGAAAGCATTTTCTTTTCTTGAACCTTAAAGTCTATTAATAAGCAGC 1620
QY 1621 ACAGATTCACATTTTATACATGAGAGATCTTCTTGTGCTGTAATACAGAGATGAC 1680
DB 1621 ACAGATTCACATTTTATACATGAGAGATCTTCTTGTGCTGTAATACAGAGATGAC 1680
QY 1681 ATCCCTTTAAAGAAAGTTTATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
DB 1681 ATCCCTTTAAAGAAAGTTTATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
QY 1741 CTTTGTAGATGATGAGATTTTGTGAGCACAATTTGGAGGCTTGAATTTGAGTGA 1800
DB 1741 CTTTGTAGATGATGAGATTTTGTGAGCACAATTTGGAGGCTTGAATTTGAGTGA 1800
QY 1801 TGGCAGGAAGAGGCCATCTCCATTTGATGATTAAGTGAACCAACTAGTCTCGGAATT 1860

Db	1801	TGGCAGAAAGGGCCATCTTCATTCATGGATGATTAAAGTGAACCAACTAGTCTCGAATT	1860
Qy	1861	CTACAGGAAGAGGGAATCAGACAGGAACCTGTGCATATMGACTTGAAGACCAAGA	1920
Db	1861	CTACAGGAAGAGGGAATCAGACTGAGGAACCTGTGACATTBGACTTGAAGACCAAGA	1920
Qy	1921	CTTTGAAATTTGCGAGCTGCTCATGTGAGTATTAATCACTGCTGTCTTTCTATTGAGT	1980
Db	1921	CTTTGAAATTTGCGAGCTGCTCATGTGAGTATTAATCACTGCTGTCTTTCTATTGAGT	1980
Qy	1981	TACAAATCTATATTTTATTTGAAGTTAAATTAAGAAAAAATTTTCAAGAAAAA	2040
Db	1981	TACAAATCTATATTTTATTTGAAGTTAAATTAAGAAAAAATTTTCAAGAAAAA	2040
Qy	2041	A 2041	
Db	2041	A 2041	

Query Match	43.9%	Score 896	DB 24	Length 2130
XX	ABN59708	standard; cDNA; 2130 BP.		
XX	ABN59708			
XX	ABN59708			
DT	28-JUN-2002	(first entry)		
XX				
DE	Novel human coding sequence SEQ ID NO: 119.			
XX				
KM	Human; antihaemic; vulnerary; antiinflammatory; immunomodulator;			
KM	antiinfectivity; cerebroprotective; cytoslatic; rheumatic; gene therapy;			
KM	neuroprotective; antiparkinsonian; protein therapy; EST;			
XX	expressed sequence tag; gene; ss.			
OS	Homo sapiens.			
XX	MO200222660-A2.			
PN				
XX	21-MAR-2002.			
PD				
XX	10-SEP-2001; 2001MO-US26015.			
PF				
XX	11-SEP-2000; 2000US-0659671.			
PR				
XX				
PA	(HYSE-) HYSEQ INC.			
XX				
PI	Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;			
P1	Xue AD, Yang Y, Wehrman T, Drmanac RT;			
XX				
DR	WPI: 2002-292408/33.			
DR	P-PSDB: ABB97295.			
XX				
FT	An isolated polynucleotide for treating diseases associated with its			
PT	encoded polypeptide such as cancer and multiple sclerosis -			
XX				
XX	Claim 1; SEQ ID NO 119; 509pp; English.			
XX				
CC	The present invention provides the protein and coding sequences of 444			
CC	novel human proteins. These were isolated from expressed sequences tags			
CC	(ESTs). They can be used to stimulate cell growth, to regulate			
CC	haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth			
CC	e.g. in burn treatment, to regulate the immune system e.g. to treat			
CC	multiple sclerosis, to regulate activin or inhibin e.g. to treat			
CC	infertility, to regulate haemostasis or thrombolysis e.g. to treat			
CC	stroke and cancer, to screen for drugs, to treat inflammatory conditions			
CC	e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.			
CC	Parkinson's disease. The present sequence is a coding sequence of the			
XX	invention.			
XX				
SD	Sequence 2130 BP; 620 A; 439 C; 522 G; 549 T; 0 other;			

Best Local Similarity 100.0%; Pred. No. 0;
Matches 896; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	44	GTGGGCAGTGGGTCTAGGGGGCTGCAGACGACACCTACTCTTTGAATGATTTCTGGGGCC	10
Db	56	GTGGGCAGTGGGTCTAGGGGGCTGCAGACGACACCTACTCTTTGAATGATTTCTGGGGCC	11
OY	104	ACGGGCACTGGCAATTCACAGCTGGGCTTGACGTTAGGGCAAGGGCTTGGCGGTGAGATC	16
Db	116	ACGGGCACTGGCAATTCACAGCTGGGCTTGACGTTAGGGCAAGGGCTTGGCGGTGAGATC	17
OY	164	GTACGGCCTTACTCCATGAGAGTCTATGAAGGCTTAGACATCATCACCACAAAGGTTTCT	22
Db	176	GTACGGCCTTACTCCATGAGAGTCTATGAAGGCTTAGACATCATCACCACAAAGGTTTCT	23
OY	224	GCCCAAGAGAGAGATTCGTGCGGACACCATGATCACCTTTGTGATTCCTCTTGATCC	28
Db	236	GCCCAAGAGAGAGATTCGTGCGGACACCATGATCACCTTTGTGATTCCTCTTGATCC	29
OY	284	AATTACACAGTGTGTGACTTCAGAAATAGACACTGCTGTGATTGAATATATTTGCC	34
Db	296	AATTACACAGTGTGTGACTTCAGAAATAGACACTGCTGTGATTGAAGATATATTTGCC	35
OY	344	CGACACAAATTCCTATGTGTGGGAGGAACCAATTTTCATTTGAATCTGCGCTCG	40
Db	356	CGACACAAATTCCTATGTGTGGGAGGAACCAATTTTCATTTGAATCTGCGCTCG	41
OY	404	AAAGTTCTTTCATATCCCAAGGCCCCAGAGATGGGCACTGTGAAGAGTATTTGACGAAAA	46
Db	416	AAAGTTCTTTCATATCCCAAGGCCCCAGAGATGGGCACTGTGAAGAGTATTTGACGAAAA	47
OY	464	GTGGAGCTTAAAGAGAGATGGTCTTGACTTTCACAAAGCCTTAACCCAGGTGGACCCA	52
Db	476	GTGGAGCTTAAAGAGAGATGGTCTTGACTTTCACAAAGCCTTAACCCAGGTGGACCCA	53
OY	524	GAATGCGCTCCCAAGCTGGATGCATGCATGACAAACCAAGGAGCCAGAGACTTGCAGATT	58
Db	536	GAATGCGCTCCCAAGCTGGATGCATGCATGACAAACCAAGGAGCCAGAGACTTGCAGATT	59
OY	584	TTTGAAGAAACAGGAATCTCTCATATGATGAATTTCTCATGTCTCAACATACGGAGAAAGT	64
Db	596	TTTGAAGAAACAGGAATCTCTCATATGATGAATTTCTCATGTCTCAACATACGGAGAAAGT	65
OY	644	GGTGGTCCCTTGGAGGTCTCTGGAAGTTCTTAAACCTTGTCAATCCTTTGGCTTCAATGC	70
Db	656	GGTGGTCCCTTGGAGGTCTCTGGAAGTTCTTAAACCTTGTCAATCCTTTGGCTTCAATGC	71
OY	704	GACCAGCAGTTTAAATGAGCGCTTGGATTAAGAGGTGATGACATGCTTCTGCTGGG	76
Db	716	GACCAGCAGTTTAAATGAGCGCTTGGATTAAGAGGTGATGACATGCTTCTGCTGGG	77
OY	764	CTCTTGAGGAACTAAGAGATTTTCAACAACCTTAATTCAGAGAGATGTTTGGGAAAT	82
Db	776	CTCTTGAGGAACTAAGAGATTTTCAACAACCTTAATTCAGAGAGATGTTTGGGAAAT	83
OY	824	AGCCAGGACTATCAACATGGTATCTTCCAATCAATTTGGCTTCAAGGAATTTCAAGAGTAC	88
Db	836	AGCCAGGACTATCAACATGGTATCTTCCAATCAATTTGGCTTCAAGGAATTTCAAGAGTAC	89
OY	884	CTGATCACTAGGAAATATGCACACAGGAGACTATTAATCAACAGCTTCTTAAGAAAGG	939
Db	896	CTGATCACTAGGAAATATGCACACAGGAGACTATTAATCAACAGCTTCTTAAGAAAGG	951

RESULT 3	
ABA06500	
ID	ABA06500 standard; cDNA; 717 BP
XX	
XX	ABA06500;
AC	
XX	10-JAN-2002 (first entry)
DT	
XX	Human cDNA SEQ ID NO: 166.
DE	

XX	Human: gene therapy; neural disorder; immune system disorder;	PR	14-SEP-2000;	2000US-233065P.
KM	muscular disorder; reproductive disorder; gastrointestinal disorder;	PR	21-SEP-2000;	2000US-234223P.
KM	pulmonary disorder; cardiovascular disorder; renal disorder;	PR	21-SEP-2000;	2000US-234274P.
KW	proliferative disorder; inflammation; ss.	PR	25-SEP-2000;	2000US-234997P.
OS		PR	25-SEP-2000;	2000US-234998P.
XX	Homo sapiens.	PR	26-SEP-2000;	2000US-235484P.
XX		PR	27-SEP-2000;	2000US-235834P.
PN	WO200154474-A2.	PR	27-SEP-2000;	2000US-235836P.
XX		PR	29-SEP-2000;	2000US-236327P.
XX		PR	29-SEP-2000;	2000US-236367P.
PD		PR	29-SEP-2000;	2000US-236368P.
XX		PR	29-SEP-2000;	2000US-236369P.
PF	17-JAN-2001; 2001WO-US01349.	PR	29-SEP-2000;	2000US-236370P.
XX		PR	02-OCT-2000;	2000US-236802P.
PR	31-JAN-2000; 2000US-179065P.	PR	02-OCT-2000;	2000US-237037P.
PR	04-FEB-2000; 2000US-160628P.	PR	02-OCT-2000;	2000US-237038P.
PR	24-FEB-2000; 2000US-164664P.	PR	02-OCT-2000;	2000US-237039P.
PR	02-MAR-2000; 2000US-163350P.	PR	02-OCT-2000;	2000US-237040P.
PR	16-MAR-2000; 2000US-169874P.	PR	13-OCT-2000;	2000US-239935P.
PR	17-MAR-2000; 2000US-190076P.	PR	13-OCT-2000;	2000US-239937P.
PR	18-APR-2000; 2000US-198123P.	PR	20-OCT-2000;	2000US-240960P.
PR	19-MAY-2000; 2000US-205515P.	PR	20-OCT-2000;	2000US-241221P.
PR	07-JUN-2000; 2000US-209467P.	PR	20-OCT-2000;	2000US-241222P.
PR	28-JUN-2000; 2000US-214886P.	PR	20-OCT-2000;	2000US-241785P.
PR	30-JUN-2000; 2000US-215135P.	PR	20-OCT-2000;	2000US-241786P.
PR	07-JUL-2000; 2000US-216647P.	PR	20-OCT-2000;	2000US-241787P.
PR	07-JUL-2000; 2000US-216880P.	PR	20-OCT-2000;	2000US-241808P.
PR	11-JUL-2000; 2000US-217487P.	PR	20-OCT-2000;	2000US-241809P.
PR	11-JUL-2000; 2000US-217496P.	PR	20-OCT-2000;	2000US-241826P.
PR	14-JUL-2000; 2000US-218290P.	PR	01-NOV-2000;	2000US-244617P.
PR	26-JUL-2000; 2000US-220963P.	PR	08-NOV-2000;	2000US-246474P.
PR	26-JUL-2000; 2000US-220964P.	PR	08-NOV-2000;	2000US-246475P.
PR	14-AUG-2000; 2000US-224518P.	PR	08-NOV-2000;	2000US-246476P.
PR	14-AUG-2000; 2000US-224519P.	PR	08-NOV-2000;	2000US-246477P.
PR	14-AUG-2000; 2000US-225213P.	PR	08-NOV-2000;	2000US-246478P.
PR	14-AUG-2000; 2000US-225214P.	PR	08-NOV-2000;	2000US-246523P.
PR	14-AUG-2000; 2000US-225266P.	PR	08-NOV-2000;	2000US-246525P.
PR	14-AUG-2000; 2000US-225267P.	PR	08-NOV-2000;	2000US-246526P.
PR	14-AUG-2000; 2000US-225268P.	PR	08-NOV-2000;	2000US-246527P.
PR	14-AUG-2000; 2000US-225270P.	PR	08-NOV-2000;	2000US-246528P.
PR	14-AUG-2000; 2000US-225447P.	PR	08-NOV-2000;	2000US-246532P.
PR	14-AUG-2000; 2000US-225757P.	PR	08-NOV-2000;	2000US-246609P.
PR	14-AUG-2000; 2000US-225758P.	PR	08-NOV-2000;	2000US-246610P.
PR	14-AUG-2000; 2000US-225759P.	PR	08-NOV-2000;	2000US-246611P.
PR	18-AUG-2000; 2000US-226279P.	PR	08-NOV-2000;	2000US-246613P.
PR	22-AUG-2000; 2000US-226681P.	PR	17-NOV-2000;	2000US-249207P.
PR	22-AUG-2000; 2000US-226868P.	PR	17-NOV-2000;	2000US-249208P.
PR	22-AUG-2000; 2000US-227182P.	PR	17-NOV-2000;	2000US-249209P.
PR	30-AUG-2000; 2000US-227009P.	PR	17-NOV-2000;	2000US-249210P.
PR	01-SEP-2000; 2000US-228924P.	PR	17-NOV-2000;	2000US-249211P.
PR	01-SEP-2000; 2000US-229287P.	PR	17-NOV-2000;	2000US-249212P.
PR	01-SEP-2000; 2000US-229343P.	PR	17-NOV-2000;	2000US-249213P.
PR	01-SEP-2000; 2000US-229344P.	PR	17-NOV-2000;	2000US-249214P.
PR	05-SEP-2000; 2000US-229345P.	PR	17-NOV-2000;	2000US-249215P.
PR	05-SEP-2000; 2000US-229509P.	PR	17-NOV-2000;	2000US-249216P.
PR	05-SEP-2000; 2000US-229513P.	PR	17-NOV-2000;	2000US-249217P.
PR	06-SEP-2000; 2000US-230437P.	PR	17-NOV-2000;	2000US-249218P.
PR	06-SEP-2000; 2000US-230438P.	PR	17-NOV-2000;	2000US-249244P.
PR	08-SEP-2000; 2000US-231242P.	PR	17-NOV-2000;	2000US-249245P.
PR	08-SEP-2000; 2000US-231243P.	PR	17-NOV-2000;	2000US-249264P.
PR	08-SEP-2000; 2000US-231244P.	PR	17-NOV-2000;	2000US-249265P.
PR	08-SEP-2000; 2000US-231413P.	PR	17-NOV-2000;	2000US-249297P.
PR	08-SEP-2000; 2000US-231414P.	PR	17-NOV-2000;	2000US-249299P.
PR	08-SEP-2000; 2000US-232080P.	PR	17-NOV-2000;	2000US-249300P.
PR	08-SEP-2000; 2000US-232081P.	PR	01-DEC-2000;	2000US-250160P.
PR	12-SEP-2000; 2000US-231968P.	PR	01-DEC-2000;	2000US-250391P.
PR	14-SEP-2000; 2000US-232397P.	PR	05-DEC-2000;	2000US-251030P.
PR	14-SEP-2000; 2000US-232398P.	PR	05-DEC-2000;	2000US-251988P.
PR	14-SEP-2000; 2000US-232399P.	PR	05-DEC-2000;	2000US-256719P.
PR	14-SEP-2000; 2000US-232400P.	PR	06-DEC-2000;	2000US-251479P.
PR	14-SEP-2000; 2000US-232401P.	PR	08-DEC-2000;	2000US-251856P.
PR	14-SEP-2000; 2000US-233063P.	PR	08-DEC-2000;	2000US-251868P.
PR	14-SEP-2000; 2000US-233064P.	PR	08-DEC-2000;	2000US-251869P.

PR 08-DEC-2000; 2000US-251989P.
PR 08-DEC-2000; 2000US-251990P.
PR 11-DEC-2000; 2000US-254097P.
PR 05-JAN-2001; 2001US-259678P.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI: 2001-476161/51.
DR P-PSDB; ABB10278.
XX
PT Isolated nucleic acid molecule encoding an inflammation-associated
PT polypeptide is used in preventing, treating or ameliorating a medical
PT condition
XX
XX Claim 1; SEQ ID NO: 166; 859pp + Sequence Listing; English.
XX
XX The present invention provides human cDNAs, proteins and related genomic
CC DNAs. These can be used in the treatment of neural, immune system,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders and inflammation. The present sequence
CC is a cDNA of the invention.
XX
XX

Sequence 717 BP; 184 A; 176 C; 187 G; 162 T; 8 other;

Query Match 23.2%; Score 473; DB 22; Length 717;

Best Local Similarity 99.7%; Pred. No. 1.3e-213; Mismatches 2; Indels 0; Gaps 0;

Matches 573; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 44 GTGGGAGTGGGCTGAGGGGCTGCAAGGACCTGCTGTAGTGAATTCGCGGGCC 103
DB 44 GTGGGAGTGGGCTGAGGGGCTGCAAGGACCTGCTGTAGTGAATTCGCGGGCC 103
QY 104 ACGGGACCGGCAATCCACGCTGGGCTTGCAGCTAGGCGGCTGCGGCTGAGATC 163
DB 104 ACGGGACCGGCAATCCACGCTGGGCTTGCAGCTAGGCGGCTGCGGCTGAGATC 163
QY 164 GTGAGCGCTGACTCCATGAGAGTCTATGAAGGCTTACATCATCACAAGTTTCT 223
DB 164 GTGAGCGCTGACTCCATGAGAGTCTATGAAGGCTTACATCATCACAAGTTTCT 223
QY 164 GTGAGCGCTGACTCCATGAGAGTCTATGAAGGCTTACATCATCACAAGTTTCT 223
DB 164 GTGAGCGCTGACTCCATGAGAGTCTATGAAGGCTTACATCATCACAAGTTTCT 223
QY 224 GCCCAAGACAGAGATTCGCGGACACATGATCAGCTTGTGTGATCCTCTTGACC 283
DB 224 GCCCAAGACAGAGATTCGCGGACACATGATCAGCTTGTGTGATCCTCTTGACC 283
QY 284 AATTACACAGTGTGACTTCAAGAAATAGAGCAACTGCTGTATGAGATATTTGCC 343
DB 284 AATTACACAGTGTGACTTCAAGAAATAGAGCAACTGCTGTATGAGATATTTGCC 343
QY 344 CGAGACAAATTCCTATTGTTGTGGAGAACCAATTATTACTTAAATCTGCTCTGG 403
DB 344 CGAGACAAATTCCTATTGTTGTGGAGAACCAATTATTACTTAAATCTGCTCTGG 403
QY 404 AAAGTCTGTCAATACAGAGCCCGAGAGATGGGACCTAGAGAAATGATGACCAAAA 463
DB 404 AAAGTCTGTCAATACAGAGCCCGAGAGATGGGACCTAGAGAAATGATGACCAAAA 463
QY 464 GTGAGACTTGAAGAGAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 523
DB 464 GTGAGACTTGAAGAGAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 523
QY 524 GAAATGCTGCCAAGTGTGATGCATGACAAACGCAAGTGGCCGAGAGCTTGCAAGTT 583
DB 524 GAAATGCTGCCAAGTGTGATGCATGACAAACGCAAGTGGCCGAGAGCTTGCAAGTT 583
QY 584 TTGGAAGAAACAGAGATCTCTCATAGTGAATTTCT 618
DB 584 TTGGAAGAAACAGAGATCTCTCATAGTGAATTTCT 618

RESULT 4
ABA06698

ID ABA06698 standard; cDNA; 717 BP.
XX
XX ABA06698;
AC
XX
XX 10-JAN-2002 (first entry)
DT
XX
XX Human cDNA SEQ ID NO: 364.
DE
XX
XX Human; gene therapy; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW proliferative disorder; inflammation; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200154474-A2.
PN
XX
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US01349.
PF
XX
XX 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 24-FEB-2000; 2000US-184664P.
PR 02-MAR-2000; 2000US-186350P.
PR 16-MAR-2000; 2000US-189874P.
PR 17-MAR-2000; 2000US-190076P.
PR 18-APR-2000; 2000US-198123P.
PR 19-MAY-2000; 2000US-205515P.
PR 07-JUN-2000; 2000US-209467P.
PR 28-JUN-2000; 2000US-214866P.
PR 30-JUN-2000; 2000US-215135P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225213P.
PR 14-AUG-2000; 2000US-225214P.
PR 14-AUG-2000; 2000US-225266P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 14-AUG-2000; 2000US-225759P.
PR 18-AUG-2000; 2000US-226279P.
PR 22-AUG-2000; 2000US-226681P.
PR 22-AUG-2000; 2000US-226688P.
PR 22-AUG-2000; 2000US-227182P.
PR 23-AUG-2000; 2000US-227009P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 06-SEP-2000; 2000US-230437P.
PR 06-SEP-2000; 2000US-230438P.
PR 08-SEP-2000; 2000US-231242P.
PR 08-SEP-2000; 2000US-231243P.
PR 08-SEP-2000; 2000US-231244P.
PR 08-SEP-2000; 2000US-231413P.
PR 08-SEP-2000; 2000US-231414P.
PR 08-SEP-2000; 2000US-232080P.
PR 08-SEP-2000; 2000US-232081P.
PR 12-SEP-2000; 2000US-231968P.

Query Match	Best Local Similarity	Score	DB	Length
Matches 573; Conservative	0; Mismatches	2; Indels	0; Gaps	0;
Sequence 717 BP; 184 A; 177 C; 187 G; 162 T; 7 other:				
0Y	44 GTGGGACAGGGGCTCAGGGGCTCGACAGGACCCCTCTTGTAGTATCTCGGGGCC	103		
Db	44 GTGGGACAGGGGCTCAGGGGCTCGACAGGACCCCTCTTGTAGTATCTCGGGGCC	103		
0Y	104 ACGGGACACGGGCAATTCACAGCTGGCGTTGACAGTACGACGCGCTCGGGGTGACATC	163		
Db	104 ACGGGACACGGGCAATTCACAGCTGGCGTTGACAGTACGACGCGCTCGGGGTGACATC	163		
0Y	164 GTACACGCTGACTCATCTCAGGTCTATATAGAGCCTACACATCATCACCACAAAGCTTCT	223		
Db	164 GTACACGCTGACTCATCTCAGGTCTATATAGAGCCTACACATCATCACCACAAAGCTTCT	223		
0Y	224 GCCCAAGAGCAGAGATCTGGCGGACACATGATCAGCTTGTGGATCTCTTGTGACC	283		
Db	224 GCCCAAGAGCAGAGATCTGGCGGACACATGATCAGCTTGTGGATCTCTTGTGACC	283		
0Y	284 AATTACACAGTGTGTGACTCAGAAATAGAGCAACTGCTGTGATTTGAAGATATATTTGCC	343		
Db	284 AATTACACAGTGTGTGACTCAGAAATAGAGCAACTGCTGTGATTTGAAGATATATTTGCC	343		
0Y	344 CGAGCAAAATTTCTATTGTTGTGGAGGAACCAATTATTCATTGAATCTCTCTCTGG	403		
Db	344 CGAGCAAAATTTCTATTGTTGTGGAGGAACCAATTATTCATTGAATCTCTCTCTGG	403		
0Y	404 AAAGTTCTTGTCAATACCAAGCCCGAGAGATGGGCACTGAGAAAGATGAGACGAAAA	463		
Db	404 AAAGTTCTTGTCAATACCAAGCCCGAGAGATGGGCACTGAGAAAGATGAGACGAAAA	463		
0Y	464 GTGAGAGCTTGAAGAGAGATGTCTTTACTTTCACAAACCCCTTAAGCCAGGTGAGCCA	523		
Db	464 GTGAGAGCTTGAAGAGAGATGTCTTTACTTTCACAAACCCCTTAAGCCAGGTGAGCCA	523		
0Y	524 GAATGCTTGCACAACTGCATCAGATGACAAAGCAAGTGGCCAGAGAGCTTCAAGTT	583		
Db	524 GAATGCTTGCACAACTGCATCAGATGACAAAGCAAGTGGCCAGAGAGCTTCAAGTT	583		

OY 584 TTTGAGAAACGAACTCTCATAGTGAATTTCT 618
|||||
DB 584 TTTGAGAAACGAACTCTCATAGTGAATTTCT 618

RESULT 5
AAS41268
ID AAS41268 standard; CDNA: 717 BP.
XX
AC AAS41268;
XX
DT 17-DEC-2001 (first entry)
XX
DE CDNA encoding novel human enzyme polypeptide #484.
XX
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.

OS Homo sapiens.
XX
PN WO20015301-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01239.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225216.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-022557.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227709.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251889.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI: 2001-465566/50.
 DR P-PSDB; AAU23398.
 XX
 PT Novel polypeptides and polynucleotides useful for diagnosing,
 PT preventing, treating neural, immune system, muscular, reproductive,
 PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous
 PT diseases
 XX
 PS Claim 4; SEQ ID NO 494; 1180bp; English.
 XX
 CC The present invention relates to the isolation of novel human enzyme
 CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
 CC encoding them. The enzyme polypeptides of the invention may comprise the
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
 CC isomerases or ligases. The sequences of the invention are useful in the
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of
 CC disorders including hyperproliferative disorders (e.g. cancer),
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders
 CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
 CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
 CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
 CC blood-related disorders (e.g. hemophilia), reproductive disorders
 CC (e.g. infertility) and infectious disorders (e.g. influenza). The
 CC polynucleotides of the invention can also be used in gene therapy.
 CC AAS40785-AAS41684 represent cDNA sequences encoding for the novel human
 CC enzyme polypeptides of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX
 SQ Sequence 717 BP; 184 A; 177 C; 187 G; 162 T; 7 other:
 Query Match 23.2%; Score 473; DB 22; Length 717;
 Best Local Similarity 99.7%; Pired. No. 1.3e-213;
 Matches 573; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 224 GCCAAGAGAGAGATCTGCCGGCACCACATGATACGCTTGTGATCCTCTTGACC 283
 Db 224 GCCAAGAGAGAGAAATCTGCCGGCACCACATGATACGCTTGTGATCCTCTTGACC 283
 QY 284 AATTACACAGTGTGACCTTGAGAAATAGACCACTCTCTGATTGAAGATATTTGGC 343
 Db 284 AATTACACAGTGTGACCTTGAGAAATAGACCACTCTCTGATTGAAGATATTTGGC 343
 QY 344 CGAGACAAATTCCTATTGTTGTGGAGAACCAATTATTACATTGAATCTCTCTG 403
 Db 344 CGAGACAAATTCCTATTGTTGTGGAGAACCAATTATTACATTGAATCTCTCTG 403
 QY 404 AAATTTCTTGATACCAAGCCCGAGAGATGGGCACTGGAAGATGATTGACCGAAA 463
 Db 404 AAATTTCTTGATACCAAGCCCGAGAGATGGGCACTGGAAGATGATTGACCGAAA 463
 QY 464 GTGAGGCTTGAAGAGAGATGCTTGTACTTACAAAGCCCTAAGCAGGTGACCA 523
 Db 464 GTGAGGCTTGAAGAGAGATGCTTGTACTTACAAAGCCCTAAGCAGGTGACCA 523
 QY 524 GAATGCTGCGCAAGCTGCATCCATGACAAACGCAAAAGTGGCCAGAGCTTGAAGT 583
 Db 524 GAATGCTGCGCAAGCTGCATCCATGACAAACGCAAAAGTGGCCAGAGCTTGAAGT 583
 QY 584 TTGAGAGAACAGGAATCTCTCATAGTGAATTTCT 618
 Db 584 TTGAGAGAACAGGAATCTCTCATAGTGAATTTCT 618
 RESULT 6
 ABL68710/c
 ID ABL68710 standard; DNA: 355 BP.
 XX
 AC ABL68710;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Kidney cancer related gene sequence SEQ ID NO:7047.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001MO-US10838.
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 22-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.

PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 28-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX (AVAL-) AVALON PHARM.
XX
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI: 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
PS Claim 1; SEQ ID 7047; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (1) of a signature gene set, where (1)
CC comprises a sequence (S) selected from 8447 sequences (given in AB61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (1) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
XX Sequence 355 BP; 112 A; 74 C; 54 G; 115 T; 0 other:

Query Match 11.8%; Score 241; DB 24; Length 355;
Best Local Similarity 99.4%; Pred. No. 9.5e-104;
Matches 341; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1690 AAGAGGTTTATGTCCTGAGCTGGCGTAAATATCATATTCAGATGCTTTGTAG 1749
DB 343 AAGAGGTTTATGTCCTGAGCTGGCGTAAATATCATATTCAGATGCTTTGTAG 284
OY 1750 ATGACTGAAGTATTTGTGAGCCACATATTGGAGTTTGTAGTTGAGTGAATGCGAGAA 1809
DB 283 ATGACTGAAGTATTTGTGAGCCACATATTGGAGTTTGTAGTTGAGTGAATGCGAGAA 224
OY 1810 AGGCGCATCTCATGAGATGATTAAGTGAACCAACTAGTCTCGGATTTTACAGAGA 1869
DB 223 AGCGCATCTCATGAGATGATTAAGTGAACCAACTAGTCTCGGATTTTACAGAGA 164
OY 1870 AGGAGGATCTGAGTGAAGAGCTGTGACATGAGACTTGAGACCAAGACTTTGAAT 1929

DB 163 AGAGGCAATCAGACTGAGGAAGCTGTGACATGAGACTTTGAAGACCAAGACTTTGAAT 104
OY 1930 TTGCAGCTGCTCATGCTGAGTATTATTCATGCTGCTTTCTTTGAGTTACAAATCT 1989
DB 103 TTGCAGCTGCTCATGCTGAGTATTATTCATGCTGCTTTCTTTGAGTTACAAATCT 44
OY 1990 ATATTTTATTCAGACTTTAAATAAGAAAAAATTTTACAGAAA 2032
DB 43 ATATTTTATTCAGACTTTAAATAAGAAAAAATTTTACAGAAA 1
RESULT 7
AAA00911
ID AAA00911 standard; cDNA; 300 BP.
XX
XX AAA00911;
AC
XX
XX 19-MAY-2000 (first entry)
DT
XX
XX Human colon cancer cell line polynucleotide sequence SEQ ID NO:902.
DE
XX
XX Human; colon cancer; tumour; diagnosis; gene expression product;
KM probe; detection; cancerous state; metastasis; identification;
KM breast cancer; oestrogen receptor-positive breast cancer; therapy;
KM oestrogen receptor-negative breast cancer; lung cancer; ss.
XX
XX Homo sapiens.
OS
XX
XX WO9958675-A2.
PN
XX
XX 18-NOV-1999.
PD
XX
XX 13-MAY-1999; 99WO-US10602.
PF
XX
XX 14-MAY-1998; 98US-0085426.
PR 15-MAY-1998; 98US-0085537.
PR 15-MAY-1998; 98US-0085696.
PR 21-OCT-1998; 98US-0105234.
PR 27-OCT-1998; 98US-0105877.
XX
XX (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lanson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX
DR WPI: 2000-126369/11.
XX
PT Polynucleotide library used to determine cancerous states of mammalian
PT cells -
XX
XX Claim 1; Page 399; 1097pp; English.
XX
XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
CC libraries constructed from human colon cancer cell lines. The present
CC invention also describes a method of detecting differentially expressed
CC genes correlated with a cancerous state of a mammalian cell, comprising
CC detecting at least one differentially expressed gene product in a test
CC sample derived from a cell suspected of being cancerous, where detection
CC of the differentially expressed gene product is correlated with a
CC cancerous state of the cell from which the test sample was derived.
CC The polynucleotides sequences can be used in a method for detecting
CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of
CC pre-metastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer.
XX

SQ Sequence 300 BP: 83 A; 69 C; 72 G; 76 T; 0 other:

Query Match 8.6%; Score 175; DB 21; Length 300;
Best Local Similarity 99.6%; Pred. No. 1.7e-72;
Matches 225; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 939 GACCTGGTCCCATGTCCTCCCTGCTATGAGCTAGAGTATGATGATGCTTCGAAGTGGG 998
|||||
Db 75 GACCTGGTCCCATGTCCTCCCTGCTATGAGCTAGAGTATGATGATGCTTCGAAGTGGG 134
|||||

QY 999 AGGAGTCTGTTCTTGAACCTGCTCTTGAATCGTCAAGATTTTCATCCAGGCGCACAGC 1058
|||||
Db 135 AAGAGTCTGTTCTTGAACCTGCTCTTGAATCGTCAAGATTTTCATCCAGGCGCACAGC 194
|||||

QY 1059 CTACAGCCACTCCCAATTAAGTGCATACATGAGCTGAGACACAGAACTTATACCC 1118
|||||
Db 195 CTACAGCCACTCCCAATTAAGTGCATACATGAGCTGAGACACAGAACTTATACCC 254
|||||

QY 1119 TGTGTGACCTCTGTGATGCAATCATCATTTGGGATCGCGCAATGGGC 1164
|||||
Db 255 TGTGTGACCTCTGTGATGCAATCATCATTTGGGATCGCGCAATGGGC 300
|||||

RESULT 8
ABA42839/c
ID ABA42839 standard; DNA: 457 BP.
XX ABA42839;
XX
XX 01-FEB-2002 (first entry)
XX
XX Human breast cell single exon nucleic acid probe #1534.
XX
XX Human: microarray; single exon probe; gene expression; breast;
XX disease; cancer; ss.
XX
XX Homo sapiens.
XX
XX W0200157271-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00662.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR:
XX
XX WPI: 2001-496933/54.
XX
XX
XX New spatially-addressable set of single exon nucleic acid probes,
XX useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes -
XX
XX
XX Claim 1; SEQ ID NO 1534; 327bp + sequence listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting
XX the probes with a collection of detectably labelled nucleic acids
XX derived from mRNA of human breast, and then measuring the label
XX bound to each probe of the microarray. The probes are useful for
XX verifying the expression of regions of genomic DNA predicted to
XX encode proteins. They are useful for gene discovery, and for
XX determining predisposition and/or prognosing breast disease. Gene

CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at http://wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 457 BP: 120 A; 102 C; 110 G; 125 T; 0 other:

Query Match 7.1%; Score 145; DB 22; Length 457;
Best Local Similarity 100.0%; Pred. No. 2.8e-58;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 AGGTCTATGAAGGCTTACATCATGACCAACAGGTTTCTGGCCAGACAGAACTT 242
|||||
Db 450 AGGTCTATGAAGGCTTACATCATGACCAACAGGTTTCTGGCCAGACAGAACTT 391
|||||

QY 243 GCCGGCACACATGATCAGCTTTGTGATCCTCTTGTGACCAATTACACAGTGTGACT 302
|||||
Db 390 GCCGGCACACATGATCAGCTTTGTGATCCTCTTGTGACCAATTACACAGTGTGACT 331
|||||

QY 303 TCAGAAATAGACGACACTGCTCTGAT 327
|||||
Db 330 TCAGAAATAGACGACACTGCTCTGAT 306
|||||

RESULT 9
ABA53262/c
ID ABA53262 standard; DNA: 457 BP.
XX ABA53262;
XX
XX 01-FEB-2002 (first entry)
XX
XX Human foetal liver single exon nucleic acid probe #1567.
XX
XX Human: foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
XX Homo sapiens.
XX
XX W0200157277-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00669.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR:
XX
XX WPI: 2001-483447/52.
XX
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human foetal liver -
XX
XX
XX Claim 1; SEQ ID NO 1567; 639bp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX foetal liver. The present sequence is a single exon nucleic acid

CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at http://wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;

Query Match 7.1%; Score 145; DB 22; Length 457;
Best Local Similarity 100.0%; Pred. No. 2.8e-58;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 183 AGGCTATGAGGCGCTAGACATCATCACCACCAAGGTTTCTGCCCAAGAGAGAAATCT 242
|||||
DB 450 AGGCTATGAGGCGCTAGACATCATCACCACCAAGGTTTCTGCCCAAGAGAGAAATCT 391
OY 243 GCCGGCACCACATGATGATCAGCTTTGGATCCTTTGTGACCAATTACACAGTGGTGA 302
|||||
DB 390 GCCGGCACCACATGATGATCAGCTTTGGATCCTTTGTGACCAATTACACAGTGGTGA 331
OY 303 TCAGAAATAGAGCAACTGCTCTGAT 327
|||||
DB 330 TCAGAAATAGAGCAACTGCTCTGAT 306

RESULT 10
ID AAK01519 standard; DNA; 457 BP.
XX
AC AAK01519;
XX
DT 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe SEQ ID NO: 1510.
DE
XX

KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.

XX
XX WO200157275-A2.
XX
XX

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483446/52.

PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
PS Example 4; SEQ ID NO: 1510; 650bp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX

SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;

Query Match 7.1%; Score 145; DB 22; Length 457;
Best Local Similarity 100.0%; Pred. No. 2.8e-58;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 183 AGGCTATGAGGCGCTAGACATCATCACCACCAAGGTTTCTGCCCAAGAGAGAAATCT 242
|||||
DB 450 AGGCTATGAGGCGCTAGACATCATCACCACCAAGGTTTCTGCCCAAGAGAGAAATCT 391
OY 243 GCCGGCACCACATGATGATCAGCTTTGGATCCTTTGTGACCAATTACACAGTGGTGA 302
|||||
DB 390 GCCGGCACCACATGATGATCAGCTTTGGATCCTTTGTGACCAATTACACAGTGGTGA 331
OY 303 TCAGAAATAGAGCAACTGCTCTGAT 327
|||||
DB 330 TCAGAAATAGAGCAACTGCTCTGAT 306

RESULT 11
ID AAK26966 standard; DNA; 457 BP.
XX
AC AAK26966;
XX
DT 06-NOV-2001 (first entry)

XX Human bone marrow expressed single exon probe SEQ ID NO: 1523.
DE
XX

KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.

XX
XX WO200157276-A2.
XX
XX

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
PS Example 4; SEQ ID NO: 1523; 658bp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX

SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;

Query Match 7.1%; Score 145; DB 22; Length 457;
Best Local Similarity 100.0%; Pred. No. 2.8e-58;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 183 AGGCTATGAGGCGCTAGACATCATCACCACCAAGGTTTCTGCCCAAGAGAGAAATCT 242

|||||
Db 450 AGGCTATGAAAGCGCTAGACATCATCCAAAGAGTTTGTGCCAAGACAGAGATCT 391
Oy 243 GCGGACACCATGATCAGCTTTGTGATCCTTTGTGACCAATATACAGTGTGACT 302
Db 390 GCGGACACCATGATCAGCTTTGTGATCCTTTGTGACCAATATACAGTGTGACT 331
Oy 303 TCAGAAATAGAGCACTGCTCTGAT 327
Db 330 TCAGAAATAGAGCACTGCTCTGAT 306

RESULT 12
AA11574/c
ID AA11574 standard; DNA: 457 BP.
XX
AC AA11574;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #1507 for gene expression analysis in human cervical cell sample.
XX
KW Probe: human; microarray; gene expression; cervical epithelial cell;
XX
OS cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN MO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000CB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488901/53.
XX
XX

Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human cervical epithelial cells -
Claim 25; SEQ ID No 1507; 487bp; English.

The present invention relates to human single exon nucleic acid probes
(SENPs). The present sequence is one such probe. The SENPs are derived
from human HeLa cells. The SENPs can be used to produce a single exon
microarray, which can be used for measuring human gene expression in a
sample derived from human cervical epithelial cells. By measuring gene
expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;

Query Match 7.1%; Score 145; DB 22; Length 457;
Best Local Similarity 100.0%; Pred. No. 2.8e-58;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 183 AGGCTATGAAAGCGCTAGACATCATCCAAAGAGTTTGTGCCAAGACAGAGATCT 242
Db 450 AGGCTATGAAAGCGCTAGACATCATCCAAAGAGTTTGTGCCAAGACAGAGATCT 391

Oy 243 GCGGACACCATGATCAGCTTTGTGATCCTTTGTGACCAATATACAGTGTGACT 302
Db 390 GCGGACACCATGATCAGCTTTGTGATCCTTTGTGACCAATATACAGTGTGACT 331
Oy 303 TCAGAAATAGAGCACTGCTCTGAT 327
Db 330 TCAGAAATAGAGCACTGCTCTGAT 306

RESULT 13
AA132862/c
ID AA132862 standard; DNA: 457 BP.
XX
AC AA132862;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #1548 used to measure gene expression in human placenta sample.
XX
KW Probe: microarray; human; placenta; antenatal diagnosis;
XX
OS genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN MO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000CB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488897/53.
XX
XX

Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human placenta -
Claim 25; SEQ ID No 1548; 654bp; English.

The present invention relates to single exon nucleic acid probes (SENPs).
The present sequence is one such probe. The probes are useful for
producing a microarray for predicting, measuring and displaying gene
expression in samples derived from human placenta. The probes are useful
for antenatal diagnosis of human genetic disorders.
XX

Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;

Query Match 7.1%; Score 145; DB 22; Length 457;
Best Local Similarity 100.0%; Pred. No. 2.8e-58;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 183 AGGCTATGAAAGCGCTAGACATCATCCAAAGAGTTTGTGCCAAGACAGAGATCT 242
Db 450 AGGCTATGAAAGCGCTAGACATCATCCAAAGAGTTTGTGCCAAGACAGAGATCT 391
Oy 243 GCGGACACCATGATCAGCTTTGTGATCCTTTGTGACCAATATACAGTGTGACT 302
Db 390 GCGGACACCATGATCAGCTTTGTGATCCTTTGTGACCAATATACAGTGTGACT 331
Oy 303 TCAGAAATAGAGCACTGCTCTGAT 327
Db 330 TCAGAAATAGAGCACTGCTCTGAT 306

Tue Apr 22 09:11:13 2003

us-09-513-151-3.oli.rng

Page 14

Search completed: April 21, 2003, 21:21:56
Job time : 485 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 21, 2003, 19:01:40 ; Search time 20 Seconds
(without alignments)
6005.218 Million cell updates/sec

Title: US-09-513-151-3
Perfect score: 661
Sequence: 1 CTGCCATAGATGCGCTCCG.....TTTACAGAGAAAAA 2041

Scoring table:

	OLIGO	Xgapop	Xgapext
	60.0	60.0	60.0
	60.0	60.0	60.0
	6.0	6.0	6.0
	Delop	6.0	Delext
		7.0	7.0

Searched: 262574 seqs, 29422922 residues

Word size: 1

Total number of hits satisfying chosen parameters: 438908

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame_n2p_model -DEV=xlh
-CGN2_1/USPROO/US09513151/runat.15042003.141247.27244/app.query.fasta.1.2183
-DB-Issued_Patents_AA -QFMT=fastan -SUFFIX=oln2p.ra1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=DITS -START=1 -END=1 -MATRIX=oln2p -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09513151 -CGN1.1.16 -runat.15042003.141247.27244 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -IONCLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -TRENDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	8	1.2	378	4	US-09-215-252-17
2	8	1.2	505	4	US-09-509-902A-11
3	8	1.2	666	4	US-09-134-001C-5465
4	8	1.2	1087	1	US-08-264-002-5
5	8	1.2	1306	4	US-08-999-774A-13
6	8	1.1	15	2	US-08-686-594-14
7	8	1.1	17	2	US-08-686-594-4
8	8	1.1	18	2	US-08-686-594-1
9	8	1.1	18	2	US-08-686-594-5
10	8	1.1	18	2	US-08-686-594-11
11	8	1.1	19	2	US-08-686-594-2
12	7	1.1	19	2	US-08-686-594-6

c 13	7	1.1	19	2	US-08-686-594-12
c 14	7	1.1	20	2	US-08-686-594-3
c 15	7	1.1	20	2	US-08-686-594-7
c 16	7	1.1	20	2	US-08-686-594-13
c 17	7	1.1	26	1	US-07-942-245-287
c 18	7	1.1	26	1	US-07-942-245-325
c 19	7	1.1	26	1	US-07-942-245-348
c 20	7	1.1	73	4	US-09-134-001C-5563
c 21	7	1.1	83	4	US-09-100-802-6
c 22	7	1.1	104	1	US-09-609-324A-8
c 23	7	1.1	104	2	US-08-920-440B-8
c 24	7	1.1	104	4	US-09-173-492-8
c 25	7	1.1	104	4	US-09-173-133-8
c 26	7	1.1	104	4	US-09-165-533-8
c 32	7	1.1	116	1	US-08-209-747-18
c 33	7	1.1	116	1	US-08-458-298-18
c 34	7	1.1	129	6	5428135-6
c 35	7	1.1	132	4	US-09-034-088A-4
c 36	7	1.1	133	4	US-09-452-229-10
c 37	7	1.1	151	2	US-08-913-477-11
c 38	7	1.1	169	4	US-09-134-001C-4614
c 39	7	1.1	177	4	US-09-643-597-165
c 40	7	1.1	184	1	US-09-609-324A-2
c 41	7	1.1	184	2	US-08-920-440B-2
c 42	7	1.1	184	4	US-09-173-492-2
c 43	7	1.1	184	4	US-09-173-133-2
c 44	7	1.1	184	4	US-09-165-533-2
c 45	7	1.1	184	5	PCR-US95-12779-2

ALIGNMENTS

RESULT 1
US-09-215-252-17
Sequence 17, Application US/09215252
Patent No. 6300487
GENERAL INFORMATION:
APPLICANT: LEUNG, David W.
APPLICANT: ADONREIL, Daniel
TITLE OF INVENTION: MAMMALIAN LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE
FILE REFERENCE: 077319/0151
CURRENT APPLICATION NUMBER: US/09/215,252
PRIOR FILING DATE: 1998-12-18
PRIOR APPLICATION NUMBER: US 08/618,651
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 17
LENGTH: 378
TYPE: PRT
ORGANISM: Homo sapiens
US-09-215-252-17

Alignment Scores:
Pred. No.: 21.9
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.21%
DB: 4
Gaps: 0

US-09-513-151-3 (1-2041) x US-09-215-252-17 (1-378)
QY 669 AGTCTCTAACCCTGCATCCTT 692
DB 334 SerSerLeuThrLeuAlaSerPhe 341

```
RESULT 2
US-09-509-902A-11
: Sequence 11, Application US/09509902A
: Patent No. 6387676
: GENERAL INFORMATION:
: APPLICANT: Virca, Duke
: APPLICANT: Bird, Timothy A.
: APPLICANT: Anderson, Dirk M.
: APPLICANT: Maiken, John S.
: TITLE OF INVENTION: Human CDNs Encoding Polypeptides Having Kinase Functions
: FILE REFERENCE: 2877-US
: CURRENT APPLICATION NUMBER: US/09/509,902A
: CURRENT FILING DATE: 1999-08-03
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 11
: LENGTH: 505
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-509-902A-11

Alignment Scores:
Pred. No.: 21          Length: 505
Score: 8.00           Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21%     Indels: 0
DB: 4                  Gaps: 0

US-09-513-151-3 (1-2041) x US-09-509-902A-11 (1-505)
OY 17 TCCGCGGCGCTGCAGCAGCAGT 40
DB 496 Serva1A1A1A1A1A1A1A1A1 503

RESULT 3
US-09-134-001C-5465
: Sequence 5465, Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
: FILE REFERENCE: GTC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: CURRENT FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 5465
: LENGTH: 666
: TYPE: PRT
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5465

Alignment Scores:
Pred. No.: 20.2        Length: 666
Score: 8.00           Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21%     Indels: 0
DB: 4                  Gaps: 0

US-09-513-151-3 (1-2041) x US-09-134-001C-5465 (1-666)
OY 95 CTCGGGCGCAGCGGCGGCGGCAA 118
DB 45 LeuG1yAla1A1A1A1A1A1A1A1 52

RESULT 4
US-08-264-002-5
```

```
: Sequence 5, Application US/08264002
: Patent No. 5559019
: GENERAL INFORMATION:
: APPLICANT: GUI, JIAN-FANG
: APPLICANT: FU, XIANG-DONG
: TITLE OF INVENTION: NOVEL PROTEIN SERINE KINASE, SRPK1
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
: STREET: 1880 Century Park East, Fifth Floor
: CITY: Los Angeles
: STATE: California
: COUNTRY: USA
: ZIP: 90067
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/264,002
: FILING DATE: 22-JUN-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: TUMARKIN PH.D., LISA A.
: REGISTRATION NUMBER: P-38,347
: REFERENCE/DOCKET NUMBER: PD3590
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619/455-5100
: TELEFAX: 619/455-5110
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1087 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: IMMEDIATE SOURCE:
: CLONE: CEMK
: FEATURE:
: NAME/KEY: Protein
: LOCATION: 1..1087
US-08-264-002-5

Alignment Scores:
Pred. No.: 18.8        Length: 1087
Score: 8.00           Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.22%     Indels: 0
DB: 1                  Gaps: 0

US-09-513-151-3 (1-2041) x US-08-264-002-5 (1-1087)
OY 1910 TCAAGTCTATGTCAAGCTTCT 1887
DB 829 SerSerPrometSerGlnLeuPro 836

RESULT 5
US-08-999-774A-13
: Sequence 13, Application US/08999774A
: Patent No. 6274312
: GENERAL INFORMATION:
: APPLICANT: Gish, Kurt C.
: APPLICANT: Seghezzi, Wolfgang
: APPLICANT: Shanahan, Frances
: APPLICANT: Lees, Emma M.
: APPLICANT: McClanahan, Terrill K.
: TITLE OF INVENTION: Intracellular Regulatory Molecules;
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DNAX Research Institute
```

STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/999, 774A
FILING DATE: 10-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US*60/032, 818
FILING DATE: 11-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34, 090
REFERENCE/DOCKET NUMBER: DX0646
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)496-1200
TELEFAX: (650)852-9196
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1306 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-999-774A-13

Alignment Scores:
Pred. No.: 18.4 Length: 1306
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.22% Indels: 0
DB: 4 Gaps: 0

US-09-513-151-3 (1-2041) x US-08-999-774A-13 (1-1306)
QY 1307 TCTGCGGCGATCCCTCCCTT 1284
|||||
DB 637 SerialaendlylleProserleu 644

RESULT 6
US-08-686-594-14
Sequence 14, Application US/08686594
Patent No. 5856127
GENERAL INFORMATION:
APPLICANT: POWELL, WILLIAM A.
APPLICANT: MAYNARD, CHARLES A.
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: P.O. BOX 1051, CLINTON SQUARE
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/686,594
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: TIMIAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 20884/50
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-686-594-14

Alignment Scores:
Pred. No.: 33.2 Length: 15
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.06% Indels: 0
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-14 (1-15)
QY 152 CGAGCGGCTGCCTAGCTGCA 132
|||||
DB 2 ArgAlaAlaGlyLeuAlaAla 8

RESULT 7
US-08-686-594-4
Sequence 4, Application US/08686594
Patent No. 5856127
GENERAL INFORMATION:
APPLICANT: POWELL, WILLIAM A.
APPLICANT: MAYNARD, CHARLES A.
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: P.O. BOX 1051, CLINTON SQUARE
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/686,594
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: TIMIAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 20884/50
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-686-594-4

Alignment Scores:
Pred. No.: 32.6 Length: 17
Score: 7.00 Matches: 7

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.06%
DB: 2
Conservative: 0
Matches: 0
Indels: 0
Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-4 (1-17)

QY 152 CGAGCCGCTGGCTAGCTGCA 132

DB 3 ArgAlaAlaGlyLeuAlaAla 9

RESULT 8

US-08-686-594-1
Sequence 1, Application US/08686594

Patent No. 5856127

GENERAL INFORMATION:

APPLICANT: POWELL, WILLIAM A.

APPLICANT: MAYNARD, CHARLES A.

TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP

STREET: P.O. BOX 1051, CLINTON SQUARE

CITY: ROCHESTER

STATE: NEW YORK

COUNTRY: USA

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/686,594

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: TIMIAN, SUSAN J.

REGISTRATION NUMBER: 34,103

REFERENCE/DOCKET NUMBER: 20884/50

TELECOMMUNICATION INFORMATION:

TELEPHONE: 716-263-1636

TELEFAX: 716-263-1600

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-686-594-1

Alignment Scores:

Pred. No.: 323

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.06%

DB: 2

US-09-513-151-3 (1-2041) x US-08-686-594-1 (1-18)

QY 152 CGAGCCGCTGGCTAGCTGCA 132

DB 4 ArgAlaAlaGlyLeuAlaAla 10

RESULT 9

US-08-686-594-5

Sequence 5, Application US/08686594

Patent No. 5856127

GENERAL INFORMATION:

APPLICANT: POWELL, WILLIAM A.

APPLICANT: MAYNARD, CHARLES A.

TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP

STREET: P.O. BOX 1051, CLINTON SQUARE

CITY: ROCHESTER

STATE: NEW YORK

COUNTRY: USA

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/686,594

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: TIMIAN, SUSAN J.

REGISTRATION NUMBER: 34,103

REFERENCE/DOCKET NUMBER: 20884/50

TELECOMMUNICATION INFORMATION:

TELEPHONE: 716-263-1636

TELEFAX: 716-263-1600

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-686-594-5

Alignment Scores:

Pred. No.: 323

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.06%

DB: 2

US-09-513-151-3 (1-2041) x US-08-686-594-5 (1-18)

QY 152 CGAGCCGCTGGCTAGCTGCA 132

DB 4 ArgAlaAlaGlyLeuAlaAla 10

RESULT 10

US-08-686-594-11

Sequence 11, Application US/08686594

Patent No. 5856127

GENERAL INFORMATION:

APPLICANT: POWELL, WILLIAM A.

APPLICANT: MAYNARD, CHARLES A.

TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP

STREET: P.O. BOX 1051, CLINTON SQUARE

CITY: ROCHESTER

STATE: NEW YORK

COUNTRY: USA

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/686,594

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: TIMIAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 20884/50
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1600
TELEFAX: 716-263-1636
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-686-594-11

Alignment Scores:
Pred. No.: 323 Length: 18
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.06% Indels: 0
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-11 (1-18)

OY 152 CGAGCCGCTGCGCTAGCTGCA 132
DB 4 ArgAlaAlaGlyLeuAlaAla 10

RESULT 11
US-08-686-594-2
Sequence 2, Application US/08686594
Patent No. 5856127
GENERAL INFORMATION:
APPLICANT: POWELL, WILLIAM A.
APPLICANT: MAYNARD, CHARLES A.
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: P.O. BOX 1051, CLINTON SQUARE
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/686,594
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: TIMIAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 20884/50
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1600
TELEFAX: 716-263-1636
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-686-594-2

Alignment Scores:
Pred. No.: 321 Length: 19

Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.06% Indels: 0
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-2 (1-19)

OY 152 CGAGCCGCTGCGCTAGCTGCA 132
DB 4 ArgAlaAlaGlyLeuAlaAla 10

RESULT 12
US-08-686-594-6
Sequence 6, Application US/08686594
Patent No. 5856127
GENERAL INFORMATION:
APPLICANT: POWELL, WILLIAM A.
APPLICANT: MAYNARD, CHARLES A.
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: P.O. BOX 1051, CLINTON SQUARE
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/686,594
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: TIMIAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 20884/50
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1600
TELEFAX: 716-263-1636
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-686-594-6

Alignment Scores:
Pred. No.: 321 Length: 19
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.06% Indels: 0
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-6 (1-19)

OY 152 CGAGCCGCTGCGCTAGCTGCA 132
DB 4 ArgAlaAlaGlyLeuAlaAla 10

RESULT 13
US-08-686-594-12
Sequence 12, Application US/08686594
Patent No. 5856127
GENERAL INFORMATION:
APPLICANT: POWELL, WILLIAM A.

APPLICANT: MAYNARD, CHARLES A.
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: P.O. BOX 1051, CLINTON SQUARE
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/686,594
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: TIMIAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 20884/50
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1600
TELEFAX: 716-263-1636
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-686-594-12

Alignment Scores:
Pred. No.: 321 Length: 19
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.06% Indels: 0
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-12 (1-19)
OY 152 CGAGCCGCTGGCTAGCTGCA 132
Db 4 ArgAlaAlaGlyLeuAlaAla 10

RESULT 14
US-08-686-594-3
Sequence 3, Application US/08686594
Patent No. 5856127
GENERAL INFORMATION:
APPLICANT: POWELL, WILLIAM A.
APPLICANT: MAYNARD, CHARLES A.
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: P.O. BOX 1051, CLINTON SQUARE
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/686,594
FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: TIMIAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 20884/50
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1600
TELEFAX: 716-263-1636
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-686-594-3

Alignment Scores:
Pred. No.: 319 Length: 20
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.06% Indels: 0
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-3 (1-20)
OY 152 CGAGCCGCTGGCTAGCTGCA 132
Db 4 ArgAlaAlaGlyLeuAlaAla 10

RESULT 15
US-08-686-594-7
Sequence 7, Application US/08686594
Patent No. 5856127
GENERAL INFORMATION:
APPLICANT: POWELL, WILLIAM A.
APPLICANT: MAYNARD, CHARLES A.
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: P.O. BOX 1051, CLINTON SQUARE
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/686,594
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: TIMIAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 20884/50
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1600
TELEFAX: 716-263-1636
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-686-594-7

Alignment Scores:

Pred. No.:	319	Length:	20
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.06%	Indels:	0
DB:	2	Gaps:	0

US-09-513-151-3 (1-2041) x US-08-686-594-7 (1-20)

QY 152 CGAGCCGCTGGCCTAGCTGCA 132

Db 4 ArgAlaAlaGlyLeuAlaAla 10

Search completed: April 21, 2003, 19:19:23
Job time : 26 secs

PD 21-MAR-2002.
 XX 10-SEP-2001: 2001WO-US26015.
 PF 11-SEP-2000: 2000US-0659671.
 PR
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 P1 Xue AD, Yang Y, Wehrman T, Drmanac RT;
 XX WPI: 2002-292408/33.
 DR N-PSDB: ABN32481.
 XX
 FT An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis -
 XX
 XX Example 2: SEQ ID NO 563: 509pp; English.
 XX
 CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate actin or inhibit e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention.
 XX
 SO Sequence 411 AA:
 Alignment Scores:
 Pred. No: 1,42e-206 Length: 411
 Score: 1989.00 Matches: 383
 Percent Similarity: 93.43% Conservative: 1
 Best Local Similarity: 93.19% Mismatches: 1
 Query Match: 55.64% Indels: 26
 DB: Gaps: 1
 US-09-513-151-3 (1-2041) x ABB97295 (1-411)
 OY 179 ATGCAGTCTATGAGGCTAGACATCATCCACCAAGGTTTCCCAAGACGAGAGA 238
 DB 1 MetGlnValIYrGluGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnIArg 20
 OY 239 ATCTCCGCGACACATGATGAGCTTTGTGATCCTCTTGACCAATTAACAGCTGCTG 298
 DB 21 IleCysArgHisHisMetIleSerPheValAspProLeuValThrAsnIleThrValVal 40
 OY 299 GACTTCAGAAATAGACAGCACTGCTGTGATGACATATATTTGCCCGAGCAAAATTCCT 358
 DB 41 AspPheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIlePro 60
 OY 359 ATTGTTGGGAGGAACCAATTATTAATCTGCAATCTGCTGGAAGTCTTCTCAT 418
 DB 61 IleValValIglYglYThrAsnIleTyrIleGlnSerLeuLeuTyrPlyValLeuValAsn 80
 OY 419 ACCAAGCCCCAGAGATGGGCACTGAGAAAGTATGACCCGAAAGTGGCTGCCAG 538
 DB 81 ThrIleProGlnGlnIleMetGlyThrGlnLysValIleAspArgLysValGlnLeuGlnLys 100
 OY 479 GAGATGGTCTTTGATTTACAAACGCCCTAAGCCAGGTGAGCCGAGAAATGGCTGCCAG 538
 DB 101 GluAspArgLysLeuValIleHisLysArgLeuSerGlnValAspProGlnMetAlaIalaLys 120
 OY 539 CTGCAATCCACATGACAAACGCAAAAGTGGCCAGAGCTTGCAAGTTTGAAGAACAAGCA 598
 DB 121 LeuHisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGlnGlnIleThrGly 140
 OY 599 ATCTCTCATAGATTTCTCCATGCTACACATAGGAGAAAGGAGGTGGTCCCTTGA 658
 DB 141 IleSerHisSerGlnPheLeuHisArgGlnHisThrGlnGlnIleGlyGlyProLeuGln 160

OY 659 GGTCCCTGTGAAGTTCTCTAACCCCTTGATCCTTTGGCTTCAATGCTGACACGAGCTTCTA 718
 DB 161 GlyProLeuLysPheSerAsnProCysIleLeuTyrPheHisAlaAspGlnIalaValLeu 180
 OY 719 GATGAGCGCTTGATTAAGAGGCTGATGACATGCTGCTGCTGGGCTTGAGAGCACTA 778
 DB 181 AspGlnArgLeuAspLysArgValAlaAspSerPheLeuAlaIleGlyLeuLeuGlnIleu 200
 OY 779 AGAGATTTTTCACAGACGCTATATATCAGAAAGTCTTCGAAATATACCAAGACTATCAA 838
 DB 201 ArgAspPheHisArgArgTyrAsnGlnLysAsnValSerGlnAsnSerGlnAspTyrGln 220
 OY 839 CATGCTATCTTCCAAATCAATGCTTCAAGCAATTTACAGAGTACCTGATCACTGAGGGA 898
 DB 221 HisGlyIlePheGlnSerIleGlyPheLysGlnPheHisGlnIleTyrLeuIleThrGlnGly 240
 OY 899 AATGCAACACGAGAGACAGTACAGACGCTTAAAGAAAGCA----- 940
 DB 241 LysCysThrLeuGlnThrSerAsnGlnLeuLeuLysGlyIleGlnAlaLeuLysGln 260
 OY 940 ----- 940
 DB 261 ValThrLysArgTyrAlaArgLysGlnAsnArgTyrValLysAsnArgPheLeuSerArg 280
 OY 941 CTTGGTCCCATTTGCCCCCTGTATGGCTTAAAGGATATGATGTCGAACTGGGAG 1000
 DB 281 ProGlyProIleValProProValTyrGlyLeuGlnValSerAspValSerLysTyrGln 300
 OY 1001 GAGTCTGTTCTTGAACGCGCTTGAATCGTGAAGTTTCATCCAGGCGCACAAAGCT 1060
 DB 301 GluSerValLeuGlnPheAlaLeuGlnIleValGlnSerPheIleGlnGlnHisLysPro 320
 OY 1061 ACAGCCACTCCCAATAAAGATGSCATACATGAAGCTGAGAACAGAAAGTATCACCTG 1120
 DB 321 ThrAlaThrProIleLysMetProTyrAsnGlnIleAsnLysArgSerTyrHisIleu 340
 OY 1121 TGTGACCTCTGTGATCGAATCATGATGGGATGCGCAATGGCGACGGCAATAAATCC 1180
 DB 341 CysAspLeuLysCysAspArgIleIleIleGlyAspArgGlnTyrPheAlaHisIleLysSer 360
 OY 1181 AAATCCCACTTGAAACCACTGAGAAAGAAAGAAAGATTTGACATGCTGCTCAACACC 1240
 DB 361 LysSerHisLysLeuAsnGlnLeuLysLysArgArgArgLeuAspSerAspAlaValAsnThr 380
 OY 1241 ATGAAAGTCAAGCTGTTTCCCACTATATACAAAGACCTAAAGGAAAGGATCCCA 1300
 DB 381 IleGlnSerGlnSerValSerProAspHisAsnLysGlnProLysGlnLysGlySerPro 400
 OY 1301 GCGCAGAAATGATCAAGAGCTGAATGCAGCCT 1333
 DB 401 GlnGlnAsnAspGlnGlnLeuLysCysSerVal 411
 RESULT 2
 ABB10278
 ID ABB10278 standard; Protein: 222 AA.
 AC ABB10278:
 XX
 XX 10-JAN-2002 (first entry)
 DT
 XX Human cDNA SEQ ID NO: 586.
 DE
 XX Human: gene therapy; neural disorder; immune system disorder;
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;
 KW pulmonary disorder; cardiovascular disorder; renal disorder;
 KW proliferative disorder; inflammation.
 OS Homo sapiens.
 XX
 XX MO200154474-A2.
 PN
 XX 02-AUG-2001.
 PD

[illegible]

PR	05-DEC-2000;	2000US-256719P.
PR	06-DEC-2000;	2000US-251479P.
PR	08-DEC-2000;	2000US-251856P.
PR	08-DEC-2000;	2000US-251868P.
PR	08-DEC-2000;	2000US-251869P.
PR	08-DEC-2000;	2000US-251989P.
PR	08-DEC-2000;	2000US-251990P.
PR	11-DEC-2000;	2000US-254097P.
PR	05-JAN-2001;	2001US-259678P.
XX	(HUMA-)	HUMAN GENOME SCI INC.
PA	Rosen CA,	Barash SC, Ruben SM;
PI	WPI:	2001-476161/51.
DR	N-PSDB:	ABAO6696.
XX	Isolated nucleic acid molecule encoding an inflammation-associated polypeptide is used in preventing, treating or ameliorating a medical condition -	
PT	Claim 11; SEQ ID NO: 784;	853pp + Sequence Listing; English.
XX	The present invention provides human cDNAs, proteins and related genomic DNAs. These can be used in the treatment of neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders and inflammation. The present sequence is a protein of the invention.	
CC	Sequence	222 AA;
SQ	Alignment Scores:	
Pred. NO.:	2,68e-107	Length: 222
Score:	1074.00	Matches: 214
Percent Similarity:	98.62%	Conservative: 1
Best Local Similarity:	98.17%	Mismatches: 3
Query Match:	30.04%	Indels: 1
DB:	22	Gaps: 0
US-09-513-151-3 (1-2041) x ABB10476 (1-222)		
OY	2 TGCCTAAGATGGCGGTCGCCGTGGCGGCTGCACAGCAACTTCTCTGGCAATGGCGCTCAGG	61
Db	1 CyshtslswmetatlaservalAlaAlaAlaArqAlaValProvalGlserGlyleuaTg	20
OY	62 GGCTGCACAGCACCTACCTCTGTAGTAATTCTCGGSCCAAGCGGCAACGAATCC	121
Db	21 GlyleuglnarqthrleuproleualValilleuonlYalathrqlYthrglyLysSer	40
OY	122 AGCGTGCGGTTCAGCTAGGCCACGGCGCTGGCGGAGTAGATCTCAGCGCTGACTCATG	181
Db	41 ThrleualaleuqInleuqIyGlnarqleuqIyGIyIulelvalserlaaaspsrmet	60
OY	182 CAGCTCATGAAAGGCTTAGACATCATACCACAAGAAGTTCTGCCAACAGCGCAAGATC	241
Db	61 GlnvallylrqIuqIyleuaspllerThrasnlysaIserlaagInguInarqIle	80
OY	242 TGCCGCGACCAATGATCAGCTTGTGTGATCTCTTGTCACCAATTACACAGTGTGGAC	301
Db	81 CyslrqhtslsmetlleseprphevalAsproleuevalThrnsntYrthralValasp	100
OY	302 TTCAGAAATAGACCAACTGCTCTGATTTGAAGATATATTGCCCGGAGCAAAATTCCTATT	361
Db	101 phearqsnahayalaatrAlaIleuIlleglAsplleheAlaIrqasplysIleproIle	120
OY	362 GTTGTGGAGCAACAATTATTTACATTGAATCTCTGCTGGAAGATTCTTGTCATATCC	421
Db	121 ValValIgLy***ThrAsnTYrTyllleglUsetleuenuTrpylsValleuValasnThr	140
OY	422 AAGCCCCAGAGATGGGCACTGAGAAAGTATTTGACCGAAAGTGAGCTTGAAAAGGAG	481
Db	141 LysrroqInguIn**GIyThrGIuLyValIIeAsparqIySValGIuIeuqInuLyGIu	160

Oy 482 GATGCTTGTACTCACAACGCGTAAAGCCAGGTGACCCAGAAATGCTGCCAGTGC 541
|||||
Db 161 AspglyLeuValIleuHisLysArgLeuSerGlnValAspProGluMetAlaIalLysLeu 180
Oy 542 CATCCACATGCAAAACGCCAAGTGCCGAGACCTTGCAAGTTTGTGAAGAAACGAATC 601
|||||
Db 181 HisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGluGlnThrGlyIle 200
Oy 602 TCTCATGAGTAATTTTTCATGCTACATAC-GGAGACAGGTGGTGCTCCCT 654
|||||
Db 201 SerHisSerGluPheLeuHisArgGlnHisThrGlyArgLysTyrTrpSerPro 218
RESULT 4
AAU23398
ID AAU23398 standard; Proteiny 222 AA.
XX
AC AAU23398:
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human enzyme polypeptide #484.
XX
KW Human: oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase: hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
KW nephrotoxic; anticoagulant.
XX
OS Homo sapiens.
XX
PN WO200155301-A2.
XX
PD 02-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US01239.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.

PR 17-NOV-2000: 2000US-0249212.
 PR 17-NOV-2000: 2000US-0249213.
 PR 17-NOV-2000: 2000US-0249214.
 PR 17-NOV-2000: 2000US-0249215.
 PR 17-NOV-2000: 2000US-0249216.
 PR 17-NOV-2000: 2000US-0249217.
 PR 17-NOV-2000: 2000US-0249218.
 PR 17-NOV-2000: 2000US-0249244.
 PR 17-NOV-2000: 2000US-0249245.
 PR 17-NOV-2000: 2000US-0249264.
 PR 17-NOV-2000: 2000US-0249265.
 PR 17-NOV-2000: 2000US-0249297.
 PR 17-NOV-2000: 2000US-0249299.
 PR 17-NOV-2000: 2000US-0249300.
 PR 01-DEC-2000: 2000US-0250160.
 PR 05-DEC-2000: 2000US-0250391.
 PR 05-DEC-2000: 2000US-0251030.
 PR 05-DEC-2000: 2000US-0251988.
 PR 06-DEC-2000: 2000US-0256719.
 PR 08-DEC-2000: 2000US-0251479.
 PR 08-DEC-2000: 2000US-0251856.
 PR 08-DEC-2000: 2000US-0251868.
 PR 08-DEC-2000: 2000US-0251869.
 PR 08-DEC-2000: 2000US-0251989.
 PR 11-DEC-2000: 2000US-0254097.
 PR 05-JAN-2001: 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM:
 DR MPI: 2001-465566/50.
 DR N-PSDB: AAS41268.
 PT Novel polypeptides and polynucleotides useful for diagnosing,
 PT preventing, treating neural, immune system, muscular, reproductive,
 PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous
 PT diseases
 XX
 XX
 PS Claim 11: SEQ ID NO 1394: 1180pp: English.
 XX
 CC The present invention relates to the isolation of novel human enzyme
 CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
 CC encoding them. The enzyme polypeptides of the invention may comprise the
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
 CC isomerases or ligases. The sequences of the invention are useful in the
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of
 CC disorders including hyperproliferative disorders (e.g. cancer),
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders
 CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
 CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
 CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
 CC blood-related disorders (e.g. haemophilia), reproductive disorders
 CC (e.g. infertility) and infectious disorders (e.g. influenza). The
 CC polynucleotides of the invention can also be used in gene therapy.
 CC AU022915-AU023814 represent the novel human enzyme polypeptides of the
 CC invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIP0
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX
 XX Sequence 222 AA:

Alignment Scores:

Pred. No.: 2,68e-107 Length: 222
 Score: 1074.00 Matches: 214
 Percent Similarity: 98.62% Conservative: 1
 Best Local Similarity: 98.17% Mismatches: 3
 Query Match: 30.04% Indels: 1
 DB: 22 Gaps: 0

US-09-513-151-3 (1-2041) x AAU23398 (1-222)

OY 2 TGCCATAGATGGCGTCCGTGGCGGCTGCAGACGACGACTTCGTGGGAGTGGGCTCAGG 61
 DB 1 CysHisLysMetIaSerValAlaAlaAlaArgAlaValProValIcylserGlyLeuArg 20
 OY 62 GGCCTGCAACGGACCTTACCTCTTGTAGTGAATTCGGGGCCAGCGGCAACGGCAATCC 121
 DB 21 GlyLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40
 OY 122 ACCGTGGCGTTCGAGTACGAGCCGCGGGCGGAGACATCGTCAGCCGTGCATTC 181
 DB 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyValLeuValSerIaAspSerMet 60
 OY 182 CAGGTCTATGAAGCGCTAGACATCATCACCACAAAGTTTCTGCCAAGACGACAGAAATC 241
 DB 61 GlnValTyrGlnClyLeuAspIleIleThrAsnLysValSerIaIcInglGlnArgIle 80
 OY 242 TGCCGGCCACCATGATCATGACCTTGTGGATCCTCTGTGACCAATTACACAGTGGTACG 301
 DB 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrThrValValAsp 100
 OY 302 TTCAGAAATAGACCACTGCTGTGATTGAAGATATATTGCCCGAGACAAATTCCTATT 361
 DB 101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle 120
 OY 362 GTTGCGGAGGACCAATTATTACATTGAATCTGCTGTGAAAGTTCTTGTGCATTACC 421
 DB 121 ValValGly***ThrAsnTyrTyrIleGluSerLeuLeuTyrLysValLeuValAsnThr 140
 OY 422 AACCCCGAGAGATGGCACCTGAGAAAGTATTGACCGAAAGTGGACGCTTGAAGAGAG 481
 DB 141 LysProGlnGlu***GlyThrGluLysValIleAspArgLysValGluLeuGlnLysGlu 160
 OY 482 GATGCTCTGTACTTACAAACCCCTAACCCAGCTGACCCAGCAAAATGCTGCCAAGCTG 541
 DB 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaAlaLysLeu 180
 OY 542 CATCCACATGACAAACGCAAGAGTGGCCAGAGCTTCAAGTTTGAAGAAACAGCAATC 601
 DB 181 HisProHisAspLysArgLysValAlaIaArgSerLeuGlnValPheGluGluThrGlyIle 200
 OY 602 TCTCATAGTGAATTTCTCCATCTGCTCAACATAC-GGAAGAAGTGCTGCTCCCT 654
 DB 201 SerHisSerGluPheLeuHisArgGlnHisThrGlyArgLysTyrPrrPserPro 218
 RESULT 5
 ID AA02526 standard: Protein: 430 AA.
 AC AA02526:
 XX
 DT 15-JUL-1999 (first entry)
 XX
 DE Amino acid sequence of the GRO-1 protein.
 XX
 KW gro-1 operon: gro-1 gene; gop-1 gene; gop-2 gene; gop-2 gene;
 KW hap-1 gene: cancer; aging; longevity; tumour formation;
 KW physiological clock.
 XX
 OS Caenorhabditis elegans.
 OS
 PN W09910482-A1.
 XX
 PD 04-MAR-1999.
 XX
 PF 20-AUG-1998: 98WO-CA00803.
 XX
 PR 25-AUG-1997: 97CA-2210251.
 XX
 PA (UYMC-) UNIV MCGILL.
 XX
 PI Barnes T, Hekimi S, Lakowski B, Lemieux J:

Score: 397.50 Matches: 98
 Percent Similarity: 54.51% Conservative: 59
 Best Local Similarity: 34.03% Mismatches: 100
 Query Match: 11.12% Indels: 31
 DB: 23 Gaps: 7

US-09-513-151-3 (1-2041) x ABB48079 (1-305)

```

OY 77 CTACCTCTTGTAGTATCTCGGGCCACGGCCGCAATTCACGCTGCGCTTGCAG 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 ILeProValIleValIleValIleProThraValIleGlyLysThraSerLeuSerIleGlu 23

OY 137 CTAGCCGACGGCTCGGGCTGAGATCGTCAAGCGCTGATCCATGCTCATGTAAGC 196
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 24 LeuAlaLysLysLeuAspGlyGluIleIleSerGlyAspSerMetGlnValIleThrGly 43

OY 197 CTAGACATCATCACCAAGAGTTTCTGCCACAGACAGAGAAATCTGCCGACCAATG 256
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44 LeuAspIleGlyThrAlaLysIleThrProGluMetAspGluIleLysHisThrLeu 63

OY 257 ATCAGCTTTGGATCCCTCTGTGACAAATACAGTGGTGGACTTCAGAAATAGACA 316
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 IleAspValIleThrAspPro---SerGluProPheThrAlaAlaLysPheGluThrGluThr 82

OY 317 ACTGCTCTGATTAAGATATATTTGCCGACAAATTTCTATTGTTGGAGAGAAC 376
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 83 ArgLysThrPleIleGluThrIleHisGlnAlaGlyLysLeuProIleIleValIleGlyThr 102

OY 377 AATTATACATTAATCTCTGCTCTGCAAGAGTTCTTGTCAATCCAGCCCCAGAGATG 436
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103 GlyLeuThrIleGlnSerValPheThrAspThrAspPheGlyAsnValSerGluAspLys 122

OY 437 GGCACATGAGAAGTATGACCGAAGTGGAGCTTGAAGAGAGAGATGCTTGTACTT 496
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 Ala-----TyrTrpAlaGluLeuGluGlnLeuAsnLysThrLeuLeu 136

OY 497 CACAAAGCCCTAAGCCAGTGGAGCCCAAGAAATGCTGCCAAGCTGACCAATGACAA 556
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 137 TrpGlnMetLeuGlnGlnGlnAspProGluSerAlaAlaGlnIleHisGluAsnLys 156

OY 557 CGCAAGTGGCCAGAGAGCTTGCAGAGTTTGAAGAAAGCAATCTTCATAGTGAATT 616
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 157 ArgAlaGlyAlaIleArgAlaLeuGlnValMetHisLeuThrGlyLysProPheSerGluThr 176

OY 617 ---CTCCATCGTCACATACGGAAGAGAGTGGTGGCTCCCTTGGAGGTCCTGAAAGTTC 673
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 GlnValHisAsnValLeuAsnAspThr----- 185

OY 674 TCTAACCTTCATCCTTTGGCTTCATGCTGACACGAGAGTTCTAGATGAGCGCTTGGAT 733
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186 TyrLysProLeuPheLeuGlyLeuAsnLysPheArgAlaLeuLeuThrGluArgIleAsn 205

OY 734 AAGAGGGTGGATGACATCTGCTGCTGCTGCTTGGAGAGAACTAAGAGATTTTCACAGA 793
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 206 GlnArgValAsnLeuMetPheGluGlnGlyLeuValThrGluAlaLys-----Lys 222

OY 794 CGCTATATACAGAAAGTTCGGAATATGACGACATCAACATGATATCTTCCAA 853
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 LeuThrAspGlnHisLeuValAspValProAla-----ValLys 235

OY 854 TCAATTTGGCTTCAAGAAATTTACAGATCACTGATCACTGAGGAAATGACACATGGAG 913
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 GlyIleGlyThrLysGluLeuPheProThrPhe-----GluGlyLysSerSerLeuGlu 253

OY 914 ACTAGTACAGAGCTTCTAAGAAA 937
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 254 GluAlaLysGluLeuIleGlnLys 261

RESULT 8
AAG82183
ID AAG82183 standard; Protein; 329 AA.
XX
AC AAG82183:

```

```

XX 03-SEP-2001 (first entry)
DT S. epidermidis open reading frame protein sequence SEQ ID NO:1460.
XX
XX S. epidermidis open reading frame protein sequence SEQ ID NO:1460.
DE
XX
XX Staphylococcus epidermidis SRI strain; infection; diagnosis;
KW vaccination; endocarditis.
XX
XX Staphylococcus epidermidis.
OS
XX
XX WO200134809-A2.
PN
XX
XX 17-MAY-2001.
PD
XX
XX 09-NOV-2000; 2000MO-US30782.
PF
XX
XX 09-NOV-1999; 99US-0164258.
PR
XX
XX (GLAX ) GLAXO GROUP LTD.
PA
XX
XX Kimmerly MJ;
PI
XX
XX WPI; 2001-316495/33.
XX
XX N-PSDB; AAH53033.
DR
XX
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis.
XX
XX Claim 18; Page 412; 2188pp; English.
PS
XX
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
XX
SQ Sequence 329 AA;
Alignment Scores:
Pred. No.: 4.88e-29 Length: 329
Score: 355.00 Matches: 90
Percent Similarity: 54.01% Conservative: 65
Best Local Similarity: 31.36% Mismatches: 106
Query Match: 9.93% Indels: 26
DB: 22 Gaps: 7
US-09-513-151-3 (1-2041) x AAG82183 (1-329)
OY 83 CTGTAGTATCTTGGGGCCACGGCCGCAATTCACGCTTGCAGCTAGGC 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6 LeuIleValIleValIleGlyProThrAlaSerGlyThrGluLeuSerIleGluValAla 25

OY 143 CAGGGCTCGGGCTGAGATGCTGACATCCATGACAGGCTATGAGAGCTTGAGC 202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26 LysLysPheAsnGlyGluIleIleSerGlyAspSerMetGlnValIleThrGlnGlyMetAsp 45

OY 203 ATCATACCAAGAGTTTCTGCCAAGAGAGAGAGATCTGCGGACCAACATGATCAGC 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db	46	IleGlyThrAlaLysValThrThrGlnGluMetGluGluLysLeuProHisTyrMetIleAsp	65
Oy	263	TTTGTGATGTCCTCTTGAGCCAAATTACACAGTGGTGACTTCAAGAAATAGACAACTGCT	322
Db	66	IleLeuProP---AspAlaSerPheSerAlaTyrGluPheLysAspArgAlaGluLys	84
Oy	323	CTGATTCAGACTAATATTGGCCCGACGACAAATATCTATTGTGTGGAGGAAACCAATTAT	382
Db	85	TyrIleLysAspIleThrArgArgGlyLysValProIleIleIleAlaGlyGlyThrGlyLeu	104
Oy	383	TACATTGAATCTGCTGCTGGAAAGTCTTGCAATATACAAAGCCCGACAGATGGGCAC	442
Db	105	TyrIleGlnSerLeuLeuTyrAsnTyrAlaPheGlnAspGlnSerIleSerGlnAspLys	124
Oy	443	GAGAAAGTGAATTGACCGAAAGTGTG---GACCTTGAAAAGAGAGATGCTTTACTTTCAC	499
Db	125	MetLysGlnValLysLeuLysLeuLysGluLeuGlnHisIleuHisAsnAsnLysLeuHis	144
Oy	500	AAACGCTTAAGCCAGGTGGACCCGAAATAGTCTGTCCCAAGTGCATGCATGCACAAAGC	559
Db	145	GluTyrLeuAlaSerPheAspLysGlnSerAlaLysAspIleHisProAsnAsnArgLys	164
Oy	560	AAAGTGGCCAGGACTTGGCAAGATTTTGAGAAACA---GGATTCCTGTATAGTGAATT	616
Db	165	ArgValLeuArgAlaIleGluTyrTyrLeuLysThrLysLysLeuSerSerArgLys	184
Oy	617	CTCCATGCTGACATACGAGGAAGAGTGTGTGTCCTCCCTTGAGGTCTCTGAATTCTCT	676
Db	185	LysValGlnGlnPheThrGlnAsn-----Tyr	193
Oy	677	AAACCTTCATCCTTTGGCTTCATGCTGACCCAGGCACTTACATGACGCTTGATAG	736
Db	194	AspThrLeuLeuIleGlyIleGluMetSerArgGluThrThrLeuTyrLeuArgIleAsnLys	213
Oy	737	AGGTGGATGATCATGCTTGTGTCGTGGGCTCTTGAGAACTAAGACTTATTCACAGACG	796
Db	214	ArgValAspIleMetLeuGlnHisGlyLeuPheAsnIleVal-----	227
Oy	797	TATATCAGAGAATGTTTGGGAAATAAGCCAGACATCAACATGATATTCTCAATCA	856
Db	228	-----GlnHisLeuValGlnGlnGlyPheGlnLysSerGlnIns---MetGlnAla	243
Oy	857	ATTGGCTTCAAGAAATTTACAGACTACCTGATCTACGTGAGGAAATGCACACTGGAGACT	916
Db	244	IleGlyTyrLysGlnLeu-----ValProValIleLysGlnLysIleSerMetGlnAsn	261
Oy	917	AGTAACACAGCTTTAAAGAAA	937
Db	262	AlaValGlnLysLeuLysGln	268
RESULT 9			
ABB93084			
ID	ABB93084 standard; Protein; 318 AA.		
XX	AC	ABB93084:	
XX	DT	31-MAY-2002 (first entry)	
XX	DE	Herbicidally active polypeptide SEQ ID NO 2295.	
XX	KW	Herbicidal; plant; agriculture; herbicide.	
XX	OS	Arabidopsis thaliana.	
XX	PN	MO300210210-A2.	
XX	PD	07-FEB-2002.	
XX	PF	28-AUG-2001; 2001MO-EP09892.	
XX	PR	28-AUG-2001; 2001MO-EP09892.	
XX	PA	(FARB) BAYER AG.	

XX	Tietjen K, Weidner M;
PI	WPI: 2002-269010/31.
DR	
XX	
XX	Identifying plant target proteins for herbicidally active compounds,
PT	comprising aligning and comparing nucleic acid or amino acid sequences
PR	from plant with nucleic acid or amino acid sequences from non-plant
PT	organisms -
XX	
PS	Claim 5; SEQ ID NO 2295; 261pp + Sequence Listing; English.
XX	
CC	The invention relates to identifying target proteins
CC	(ABB90790-ABB94016) for herbicidally active compounds, comprising
CC	aligning and comparing nucleic acid or amino acid sequences from plant
CC	with nucleic acid or amino acid sequences from non-plant organisms usin
CC	suitable search parameters, where plant sequences having an E-value
CC	greater by a factor of 3 than the E-value of most similar non-plant
CC	sequences are selected. The polypeptides or nucleic acids encoding them
CC	are useful for identifying modulators. The identified modulators are
CC	useful as herbicides.
XX	
SQ	Sequence 318 AA;
	Alignment Scores:
Pred. No.:	3,36e-27 Length: 318
Score:	338.00 Matches: 97
Percent Similarity:	42.23% Conservative: 58
Best Local Similarity:	Mismatches: 26.43%
Query Match:	9.45% Indels: 122
DB:	Gaps: 10
US-09-513-151-3 (1-2041) x ABB93084 (1-318)	
QY	83 CTTGTAGATGATTCTCGGGGCCACGGGCACAATTCACGTCGCGTTCAGCTAGGC 142
DB	7 MetValValIleMetCdlValIthrThrIlySerIySseSerLeuSerValAspIeuAla 26
QY	143 CAGCGGCTCGCGGGTGAGATCGTCAGCGCTGACTCATCAGCTGTATGAAGCCCTAGAC 202
DB	27 LeuHisPheIylAlaGlutIleLeuSsnSerAspIysMetGlnPheTyAspIlyLeuLys 46
QY	203 ATCAATCACCAAGAAGTTTCTGCCAAGACAGAAATCTGCCGACCACATGATCAGC 262
DB	47 IlerThrThrAsnGlnSerThrIleGluAspArgIglValProHisIleuLeuGly 66
QY	263 TTGTGGATCCCTCTGTGACCAATTAACAGAGTGGGACTTCAGAAATAAGCAACTGCT 322
DB	67 GluLeuAsnPrpGluAlaIglYglValThrlAlaIalagluPrlelryValmetAlaIalaglu 86
QY	323 CTGATTGAAGATATATTGCCCGAGACAAMAATTCCTATTGTGTGGGAGAACCAATTAT 382
DB	87 AlalleSerGluIleThrGlnArgIylsLysLeuProIleLeuAlaIglYglYserAsnSer 106
QY	383 TACATGGAATCTCTGCTCTGGAAAGTTCTTGCAATACCMAAGCCCCAGAGATGGGCACT 442
DB	107 TyrIleHisAlaIleuLeuAlaLys----- 114
QY	443 GAGAAAGTAGTATGACCGAAAAGTGGAGCGTGAAGGAAGATGCTTGTACTTCACAA 502
DB	114 ----- 114
QY	503 CGCCTAAGCCAGGTGACCCAGAAATGGCTGCCAAGTCATGCATGCATGACAAAGCAAA 562
DB	115 -----SerTyrAspProGluAsn-----TyrProPheSerAspHisLys 127
QY	563 GTGGCCAGAGACTTGCAAGTTTTTGAAGACAAGCAAACTCTCATAGATGATTTTCCAT 622
DB	127 ----- 127
QY	623 CGTCAACAATAGGAGAAGGTGGTGCCCTTGAGAGCGCTCAAGATTCCTAACCT 682
DB	128 -----GlySerIleCysSerGluLeuLysTyr--AspLys 138

```

OY 683 TGCATCCCTTGCTCATGCTGACACGAGCTGCTAGATGACCGCTTGATAGAGGCTG 742
Db 139 CysheilerpilleaspvalaspilservalleupegiluryleuSerleuargleu 158
OY 743 GATGACATGCTTGCTGCTGCTGCTTGAGGAGACTAGAGATTTTCACAGACGCTATAT 802
Db 159 Aspleuemetlyssecglymerpheclicluileaglupehlsarg----- 175
OY 803 CAGAGAATGTTTGGGAAATAAGCCAGACTATACATGCTATCTTCCATCAATTCGCC 862
Db 176 -----SerLysLysAlaProLysGluProLeuGlyIleIleTPlysAlaIleGly 191
OY 863 TTCAAGGAATTTACAGACTACTG----- 886
Db 192 ValGlnGluPheaspPtyrLeuLysMetTyrIleTPaspasnaspMetaspLysTrp 211
OY 887 -----ATCAGTGAAGGAAATGC 904
Db 212 AsprometarglysgluAlatyrcIuLysAlaValaArgAlaIleLysGluasnThrphe 231
OY 905 ACAGTGAAGACTAGTACACCACTCTTAAG-----AAGGACCTGGTCCCAAT 952
Db 232 GlnLeu---ThrLysaspGlnIleThrLysIleasnLysleuArgasnAlaGlyTrpasp 250
OY 953 GTCCCCCTGTC-----TATGGC 970
Db 251 IleLysLysValaspAlaThrAlaSerPheargGluAlaIleargAlaAlaLysGluGly 270
OY 971 TTAGAGTATGATGATGCTCGAG-----TGGAGAGAGCTGCTTCTTGAAGCTGCTCT 1024
Db 271 GluGlyValAlaIleGluMetGlnArgLysIleTrpaspasnLysGluValLeuGluProCysVal 290
OY 1025 GAATCGTGCAGAACTTTCATC 1045
Db 291 LysIleValArgSerHisLeu 297
RESULT 10
AAB88507
ID AAB88507 standard; Protein: 311 AA.
AC AAB88507;
DT 04-JUN-2001 (first entry)
DE Haemophilus influenzae essential bacterial protein SEQ ID NO:32.
XX Haemophilus influenzae; essential bacterial gene; identification;
KW otitis media; meningitis; upper respiratory tract infection;
KM infection; antimicrobial.
XX Haemophilus influenzae.
OS
PN MO200111033-A2.
PD 15-FEB-2001.
PF 03-AUG-2000; 2000WO-US21176.
PR 04-AUG-1999; 99US-0368382.
XX (ABBO ) ABBOTT LAB.
P1 Chovan LE, Hessler PE, Reich KA;
XX WPI: 2001-147511/15.
DR N-PSDB: AAF94360.
XX Essential bacterial genes from Haemophilus influenzae and methods for
PT identifying 'essential' genes that may be potential therapeutic targets
XX
PS Claim 9; Page 75-76; 185pp; English.

```

```

XX AAF94345 to AAF94409 represent essential bacterial genes from
CC Haemophilus influenzae, which encode the proteins given in AAB88492 to
CC AAB88507. The present invention also describes methods for identifying
CC essential bacterial genes (i.e. those essential to the survival of a
CC bacterium) using a transposition system. The methods are used to
CC identify essential genes from bacteria, especially H. influenzae (which
CC causes otitis media, meningitis and upper respiratory tract infections)
CC which may be used as targets for potential antimicrobial agents.
CC AAF94410 to AAF94416 represent PCR primers used in the exemplification
CC of the present invention.
XX
SQ Sequence 311 AA;
Alignment Scores:
Pred. No: 7,96e-27 Length: 311
Score: 334.50 Matches: 93
Percent Similarity: 51.52% Conservative: 60
Best Local Similarity: 31.31% Mismatches: 103
Query Match: 9,36% Indels: 41
DB: Gaps: 11
US-09-513-151-3 (1-2041) x AAB88507 (1-311)
OY 86 GTAGTGAATCTCGGCGCCACGCGACCGCAATCCAGCTGCGCTGACAGTACGCCAG 145
Db 6 IllePheLeuMetGlyProThrAlaSerGlyLysThrspLeuAlaIleGlnLeuArgSer 25
OY 146 CGGCTCGCGCGGTGAGATCGTCAGCGCTGACTCCATGCGAGTCTATGAAGCCTGACATC 205
Db 26 GlnLeuProValaGluValIleSerValaSerPheAlaLeuIleLysGlyMetaspIle 45
OY 206 ATCAACCAAGAGTTTCTGCCCAAGACAGAGAATTCGCCGCCACCATGATCAGCTTT 265
Db 46 GlyThrAlaLysProSerLysGluGluLeuAlaProHisArgLeuIleaspIle 65
OY 266 GTGAGTCTCTTGACCAATTAACAGAGTGGTGAAGTTCGAGTTCAGCAAGTGCCTGCTG 325
Db 66 LeuaspPro---SerGluSerTyrSerAlaMetasnPheargaspAlaLeuArgGlu 84
OY 326 ATGAGATATATATTGCGCCGAGACAAATTCATGTTGTGGAGAGAACCAATATATAC 385
Db 85 MetaLaspIleThrAlaGlnGlyLysIleProLeuLeuValaGlyGlyThrMetLeuTyr 104
OY 386 ATGAAATCTCGCTCTGGAAGTTCTGTCAATCCAGCCCGAGAGATGGGCACATGAG 445
Db 105 TyrLysAlaLeu-----IleGluGlyLeuSerProLeuProSerAlaaspGlu 120
OY 446 AAGTGAATTGACCGAAAGTGCAGCTGGA-----AAGGAGATGCTCTGTA 493
Db 121 AsnIle-----ArgAlaGluLeuGlnGlnLysAlaAlaGlnGlnGlyTrpAlaAla 137
OY 494 CTTGCAACAGCCCTTACAGCAGGTGAGCCAGAAATGCGTCCAGTGCATCCACATGAC 553
Db 138 LeuHisThrGluLeuAlaLysIleaspProIleSerAlaAlaArgIleasnProSerasp 157
OY 554 AAGCGCAAGTGGCCAGAGAGCTTCAAGTTTGAACAAACAGAGATCTCATAGTGA 613
Db 158 SerGlnArgIleasnArgAlaLeuGluValaPheTyrIleThrGlyLysSerLeuThrGlu 177
OY 614 TTTCTCATGCTCAACATAGGAGAGAGGTGGTGCCTCCCTGGAGGTCTCGTAAGTTC 673
Db 178 Leu-----ThrGluGluLysGlyGluAlaLeu-----ProTyraspPhe 190
OY 674 TCTAACCTTGATCCTTGTGCTTCAATGATGACGAGCAGCTGTAGATGAGCGCTTGAT 733
Db 191 ValGlnPheAlaIle-----AlaProGlnaspArgHisValaLeuHisGluArgIleGlu 208
OY 734 AAGAGGCTGATGACATGCTTGTGCTGCGCTCTTGAGAGAACTAAGAGATTTTCACAGA 793
Db 209 GlnArgPheHisLysMetIleGluLeuGlyPheGlnAlaGluValaGluLysLeuTyrAla 228
OY 794 CGTATATATGAGAGAATGTTTGGAAATAAGCCAGAGACTATGACATGCTATCTTCCA 853

```

```

Db      229 Arg-----GlyAspLeuAsnIleAsnLeuProSer-----IleArg 240
      854 TCATTTGGCTTCAGAGAAATTCACGAGTACTG----- 886
      241 CysValGlyTyrArgIleMetTrpGluTyrLeuGlnGlyAspTyrAlaTyrGluMet 260
      887 ATCACTGAGGAAATGACACACTGAGACTGATACACGCTTCTTAAGAA 937
      261 IlePheArgGlyIleCys-----AlaThrArgGlnLeuAlaLysArg 274

RESULT 11
ID      AAU91438 standard; Protein; 311 AA.
XX
AC      AAU91438;
XX
DT      18-JUN-2002 (first entry)
XX
DE      Haemophilus influenzae essential gene #16.
XX
KW      Essential bacterial gene; antifungal agent; antibacterial agent;
KW      antiparasitic agent; insecticidal agent; microbial infection;
KW      mucous membrane infection; otitis media; sinusitis; bronchitis;
KW      alveolitis; conjunctivitis; pneumonia; meningitis; epiglottitis;
KW      cellulitis; septic arthritis.
XX
OS      Haemophilus influenzae.
XX
PN      WO200218601-A2.
XX
PD      07-MAR-2002.
XX
PF      22-AUG-2001; 2001WO-US26245.
XX
PR      25-AUG-2000; 2000US-0649145.
XX
PA      (ABBO ) ABBOTT LAB.
XX
PI      Chovan LE, Hessler PE, Reich KA;
XX
DR      WPI: 2002-304258/34.
XX
N-PSDB: ABK64924.
XX
PT      Essential bacterial genes in Haemophilus influenzae necessary for
PT      bacterium's growth and survival, useful for screening inhibitors of
PT      polypeptides and developing therapeutic agents e.g. antimicrobial
XX
PS      Claim 9; Page 75-76; 185pp; English.
XX
XX
CC      The invention describes an essential bacterial gene (I) comprising a
CC      purified polynucleotide isolated from Haemophilus influenzae where (I)
CC      is essential to H. influenzae survival. The encoded polypeptide (II) is
CC      useful for screening substances that function to inhibit essential H.
CC      influenzae polypeptides by contacting (II) with the desired substances
CC      and measuring the response by a screen from specific, enzyme, general,
CC      affinity, phenotypic and binding screen. (I) and (II) are useful in
CC      developing therapeutic agents such as antifungal, antibacterial and
CC      antiparasitic agent, insecticidal agent, and preventive antimicrobial
CC      agents which are effective in preventing microbial infection or useful
CC      in treatment of that particular infection. (I) and (II) may also be
CC      useful in treatment of mucous membrane infections such as otitis media,
CC      sinusitis, bronchitis, alveolitis, conjunctivitis, pneumonia, meningitis,
CC      epiglottitis, cellulitis and septic arthritis. This is the amino acid
CC      sequence of an essential H. influenzae gene, described in the invention.
XX
SQ      Sequence 311 AA:

```

Alignment Scores: 7.96e-27 Length: 311
 Score: 334.50 Matches: 93
 Percent Similarity: 51.52% Conservative: 60
 Best Local Similarity: 31.31% Mismatches: 103

```

Query Match: 9.36% Indels: 41
DB: 23 Gaps: 11
US-09-513-151-3 (1-2041) x AAU91438 (1-311)
      86 GTAAGATATTCGGGGCCACGGCCAGCCGAAATTCACGCTGCTTCAGCTAGCCAG 145
      6 IlePheLeuMetGlyProThrAlaSerGlyLysThrAspLeuAlaIleGlnLeuArgSer 25
      146 CGGCTGGGGCGGTAGATCGTCACGCTGACTCCATGACGCTCTATGAGCCCTAGACATC 205
      26 GlnLeuProValGluValIleSerValAspSerAlaLeuLysThrLysGlyMetAspIle 45
      206 ATCCACACAGGTTTCTGCGCCAGAGCAGAGAAATGCGCGGACACATGATGACCTT 265
      46 GlyThrAlaLysProSerLysGluGluLeuAlaLeuAlaProHisArgLeuIleAspIle 65
      266 GTGATCTCTTGTGACCAATTCACAGTGTGAGCTTCAGAAATGAGCACTGCTCTG 325
      66 LeuAspPro---SerGluSerTyrSerAlaMetAsnPheArgAspAspAlaLeuArgGlu 84
      326 ATGAAGATATATTGCGCCGACGAAATTCCTATTGTGTGGAGAGAACCATTTATTC 385
      85 MetAlaAspIleThrAlaGlnGlyLysIleProLeuValGlyLysThrMetLeuTyr 104
      386 ATGAATCTCTGCTGGAAGATCTTGTCAATACCAAGCCCGAGAGATGGCCTGAG 445
      105 TyrLysAlaLeu-----IleGluGlyLeuSerProLeuProSerAlaAspGlu 120
      446 AAGATGATTGACCGAAAGTGAGCTTGA-----AAGAGATGCTTGTGA 493
      121 AsnIle-----ArgAlaGluLeuGlnGlnIleAlaIleGlnGlnIleTyrAlaAla 137
      494 CTCACAAACGCTTAAGCCGAGTGAGCCGAGAAATGCTTCCCAAGCTGCATCCACATGAC 553
      138 LeuHisThrGluLeuAlaLysIleAspProIleSerAlaIleArgIleAsnProSerAsp 157
      554 AACGCCAAGTGGCCAGAGCTTGCAGAGTTTGAAGAAACAGCAATCTCTCTAGTGA 613
      158 SerGlnArgIleAsnArgAlaLeuGluValPheTyrIleThrGlyLysSerLeuThrGlu 177
      614 TTCTCTCAGTCGACATACGAGAGAGAGTGTGTCCTTCCCTGAGCTCTGAAATTC 673
      178 Leu-----ThrGluGluLysGlyAlaAlaLeu-----ProTyrAspPhe 190
      674 TCATACCTTGATCTCTTGGCTTCATGCGACGAGGAGCTTATAGATGAGCTTGAT 733
      191 ValGlnPheAlaIle-----AlaProGlnAspArgHisValLeuHisGluArgIleGlu 208
      734 AAGAGGTGATGACATGCTTCTGCTGCGCTTGTGAGAGCACTTAAGATTTTCACAGA 793
      209 GlnArgPheHisLysMetIleGluLeuGlyPheGlnAlaGluValGluLysLeuTyrAla 228
      794 CGCTAATATTCAGAAATGTTTGGGAAATAGCCGAGACTATACATGATGATCTTCCAA 853
      229 Arg---GlyAspLeuAsnIleAsnLeuProSer-----IleArg 240
      854 TCATTTGGCTTCAGAGAAATTCACGAGTACTG----- 886
      241 CysValGlyTyrArgIleMetTrpGluTyrLeuGlnGlyAspTyrAlaTyrGluMet 260
      887 ATCACTGAGGAAATGACACACTGAGACTGATACACGCTTCTTAAGAA 937
      261 IlePheArgGlyIleCys-----AlaThrArgGlnLeuAlaLysArg 274

RESULT 12
AAC49765
ID      AAG49765 standard; Protein; 329 AA.
XX
AC      AAG49765;
XX
DT      18-OCT-2000 (first entry)
XX

```

DE	Arabidopsis thaliana protein fragment SEQ ID NO: 62990.	PR 22-JUN-1999;	99US-0138899.
XX	Protein identification: signal transduction pathway; metabolic pathway;	PR 23-JUN-1999;	99US-0140353.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	PR 23-JUN-1999;	99US-0140354.
KW	termination sequence.	PR 24-JUN-1999;	99US-0140695.
XX	Arabidopsis thaliana.	PR 28-JUN-1999;	99US-0140823.
OS	Arabidopsis thaliana.	PR 29-JUN-1999;	99US-0140991.
XX	Epil033405-A2.	PR 30-JUN-1999;	99US-0141287.
PN	06-SEP-2000.	PR 01-JUL-1999;	99US-0142154.
XX		PR 02-JUL-1999;	99US-0142055.
PD		PR 06-JUL-1999;	99US-0142390.
XX		PR 08-JUL-1999;	99US-0142803.
PF	25-FEB-2000; 2000EP-0301439.	PR 09-JUL-1999;	99US-0142920.
XX		PR 12-JUL-1999;	99US-0142977.
PR	25-FEB-1999; 99US-0121825.	PR 13-JUL-1999;	99US-0143542.
PR	05-MAR-1999; 99US-0123160.	PR 14-JUL-1999;	99US-0143624.
PR	09-MAR-1999; 99US-0123548.	PR 15-JUL-1999;	99US-0144005.
PR	23-MAR-1999; 99US-0125788.	PR 16-JUL-1999;	99US-0144085.
PR	25-MAR-1999; 99US-0126264.	PR 19-JUL-1999;	99US-0144086.
PR	29-MAR-1999; 99US-0126785.	PR 19-JUL-1999;	99US-0144325.
PR	01-APR-1999; 99US-0127462.	PR 19-JUL-1999;	99US-0144331.
PR	06-APR-1999; 99US-0128234.	PR 19-JUL-1999;	99US-0144332.
PR	08-APR-1999; 99US-0128714.	PR 19-JUL-1999;	99US-0144333.
PR	16-APR-1999; 99US-0129845.	PR 19-JUL-1999;	99US-0144334.
PR	19-APR-1999; 99US-0130077.	PR 19-JUL-1999;	99US-0144335.
PR	21-APR-1999; 99US-0130449.	PR 20-JUL-1999;	99US-0144352.
PR	23-APR-1999; 99US-0130510.	PR 20-JUL-1999;	99US-0144632.
PR	23-APR-1999; 99US-0130891.	PR 20-JUL-1999;	99US-0144884.
PR	28-APR-1999; 99US-0131449.	PR 21-JUL-1999;	99US-0144814.
PR	30-APR-1999; 99US-0132048.	PR 21-JUL-1999;	99US-0145088.
PR	30-APR-1999; 99US-0132407.	PR 21-JUL-1999;	99US-0145088.
PR	04-MAY-1999; 99US-0132484.	PR 22-JUL-1999;	99US-0145087.
PR	05-MAY-1999; 99US-0132485.	PR 22-JUL-1999;	99US-0145089.
PR	06-MAY-1999; 99US-0132486.	PR 22-JUL-1999;	99US-0145192.
PR	06-MAY-1999; 99US-0132487.	PR 23-JUL-1999;	99US-0145145.
PR	07-MAY-1999; 99US-0132863.	PR 23-JUL-1999;	99US-0145218.
PR	11-MAY-1999; 99US-0134256.	PR 23-JUL-1999;	99US-0145224.
PR	14-MAY-1999; 99US-0134218.	PR 23-JUL-1999;	99US-0145226.
PR	14-MAY-1999; 99US-0134219.	PR 26-JUL-1999;	99US-0145918.
PR	14-MAY-1999; 99US-0134321.	PR 27-JUL-1999;	99US-0145919.
PR	14-MAY-1999; 99US-0134370.	PR 27-JUL-1999;	99US-0145951.
PR	18-MAY-1999; 99US-0134768.	PR 28-JUL-1999;	99US-0146386.
PR	19-MAY-1999; 99US-0134941.	PR 02-AUG-1999;	99US-0146388.
PR	20-MAY-1999; 99US-0135124.	PR 02-AUG-1999;	99US-0146389.
PR	21-MAY-1999; 99US-0135353.	PR 03-AUG-1999;	99US-0147038.
PR	24-MAY-1999; 99US-0135629.	PR 03-AUG-1999;	99US-0147302.
PR	25-MAY-1999; 99US-0136021.	PR 04-AUG-1999;	99US-0147304.
PR	27-MAY-1999; 99US-0136352.	PR 04-AUG-1999;	99US-0147320.
PR	28-MAY-1999; 99US-0136782.	PR 05-AUG-1999;	99US-0147192.
PR	01-JUN-1999; 99US-0137222.	PR 05-AUG-1999;	99US-0147260.
PR	03-JUN-1999; 99US-0137528.	PR 06-AUG-1999;	99US-0147303.
PR	04-JUN-1999; 99US-0137502.	PR 06-AUG-1999;	99US-0147416.
PR	07-JUN-1999; 99US-0137724.	PR 09-AUG-1999;	99US-0147493.
PR	08-JUN-1999; 99US-0138094.	PR 10-AUG-1999;	99US-0147935.
PR	10-JUN-1999; 99US-0138540.	PR 10-AUG-1999;	99US-0148171.
PR	10-JUN-1999; 99US-0138847.	PR 11-AUG-1999;	99US-0148341.
PR	14-JUN-1999; 99US-0139119.	PR 12-AUG-1999;	99US-0148341.
PR	16-JUN-1999; 99US-0139452.	PR 13-AUG-1999;	99US-0148565.
PR	16-JUN-1999; 99US-0139453.	PR 16-AUG-1999;	99US-0148684.
PR	18-JUN-1999; 99US-0139454.	PR 17-AUG-1999;	99US-0149368.
PR	18-JUN-1999; 99US-0139455.	PR 17-AUG-1999;	99US-0149175.
PR	18-JUN-1999; 99US-0139456.	PR 20-AUG-1999;	99US-0149426.
PR	18-JUN-1999; 99US-0139457.	PR 20-AUG-1999;	99US-0149723.</

PR 30-AUG-1999: 99US-0151303.
 PR 31-AUG-1999: 99US-0151438.
 PR 01-SEP-1999: 99US-0151530.
 PR 07-SEP-1999: 99US-0152363.
 PR 10-SEP-1999: 99US-0153070.
 PR 13-SEP-1999: 99US-0153758.
 PR 15-SEP-1999: 99US-0154018.
 PR 16-SEP-1999: 99US-0154039.
 PR 20-SEP-1999: 99US-0154779.
 PR 22-SEP-1999: 99US-0155139.
 PR 23-SEP-1999: 99US-0155186.
 PR 24-SEP-1999: 99US-0155659.
 PR 28-SEP-1999: 99US-0156458.
 PR 29-SEP-1999: 99US-0156596.
 PR 04-OCT-1999: 99US-0157117.
 PR 05-OCT-1999: 99US-0157753.
 PR 06-OCT-1999: 99US-0157865.
 PR 07-OCT-1999: 99US-0158029.
 PR 08-OCT-1999: 99US-0158232.
 PR 12-OCT-1999: 99US-0158369.
 PR 13-OCT-1999: 99US-0159293.
 PR 13-OCT-1999: 99US-0158294.
 PR 13-OCT-1999: 99US-0159295.
 PR 14-OCT-1999: 99US-0159329.
 PR 14-OCT-1999: 99US-0159330.
 PR 14-OCT-1999: 99US-0159331.
 PR 14-OCT-1999: 99US-0159637.
 PR 14-OCT-1999: 99US-0159638.
 PR 18-OCT-1999: 99US-0159584.
 PR 21-OCT-1999: 99US-0160741.
 PR 21-OCT-1999: 99US-0160767.
 PR 21-OCT-1999: 99US-0160768.
 PR 21-OCT-1999: 99US-0160770.
 PR 21-OCT-1999: 99US-0160814.
 PR 21-OCT-1999: 99US-0160815.
 PR 22-OCT-1999: 99US-0160980.
 PR 22-OCT-1999: 99US-0160981.
 PR 22-OCT-1999: 99US-0160989.
 PR 25-OCT-1999: 99US-0161404.
 PR 25-OCT-1999: 99US-0161405.
 PR 25-OCT-1999: 99US-0161406.
 PR 26-OCT-1999: 99US-0161359.
 PR 26-OCT-1999: 99US-0161360.
 PR 26-OCT-1999: 99US-0161361.
 PR 28-OCT-1999: 99US-0161920.
 PR 28-OCT-1999: 99US-0161992.
 PR 28-OCT-1999: 99US-0161993.
 PR 29-OCT-1999: 99US-0162142.

Alignment Scores:

Pred. No.: 1,06e-26 Length: 329
 Score: 333.50 Matches: 99
 Percent Similarity: 38.948 Conservative: 56
 Best Local Similarity: 24.87% Mismatches: 105
 Query Match: 9.33% Indels: 138
 DB: 21 Gaps: 9

US-09-513-151-3 (1-2041) x AAG49765 (1-329)

QY 83 CTGTAGATATTCGCGGCGCACCGGCAAAATCCAGCTGCGCTTGACGTAGCC 142
 Db 36 ValIlePheValMetCylValThrGlySerGlyLysSerArgLeuAlaIleAspLeuAla 55
 QY 143 CACGGCTCGCGGTGATGATGTCACCGCTGATCATGCGAGGTGTATGAGCCGTAGAC 202
 Db 56 ThrArgPheGlnIleGlyIleLeuIleAsnSerAspLysIleGlnLeuIleGlyLeuAsp 75
 QY 203 ATCATACCAACAGAGTTCTGCCCAAGACAGCAATCTGCCGCGACCAATGATGACG 262
 Db 76 ValLeuThrAsnLysValThrProLysGluCysArgGlyValProIleHisLeuLeuGly 95
 QY 263 TTGTGTGATTCCTCTGTGACCAATTAACAGTGTGATGCTGCAAAATGAGCAACTGCT 322

Db 96 ValIleAspSerGluAlaGlyAsnLeuThrAlaThrGlnTyrSerArgLeuAlaSerGln 115
 QY 323 CTGATTTGAAGATATATTTGCGCCGACGAAATTCCTATGTGTGTGGAGCAATAT 382
 Db 116 AlaIleSerLysLeuSerAlaAsnLysLeuProIleValAlaGlyLysSerAsnSer 135
 QY 383 TACATTAATCTCTCTCTGTGAAGTCTTGTCAATACCAAGCCAGAGATGGCACT 442
 Db 136 TyrIleGluAlaLeu----- 140
 QY 443 GAGAAAGTATTTGACCGAAGAAAGTGGAGCTTGAAAGAGAGATGCTTGTACTTCACAA 502
 Db 140 ----- 140
 QY 503 CCGCTAAGCCAGGTGAGCCGACCAAAATGGCTGCCAAGCTGCATCAGATGACAAAGCAA 562
 Db 140 ----- 140
 QY 563 GTGGCCAGAGCTTGCAGATTTTGAGAAACAGAAATCTCAATAGTAATTTCTCCAT 622
 Db 141 -----ValAsnHisSer----- 145
 QY 623 CGTCAACATACGGAAGAGCTGGTGGCTTGGAGGCTCTGTCAAGTCTCTAACCT 682
 Db 146 -----GlyPheLeuLeuAsnTyrAspCys 154
 QY 683 TGCATCTTTGGCTTCATGCTGACACGAGCTTCTGTGATGAGCGCTGTGATGAGGGCTG 742
 Db 155 CysPheIleThrPylAspValSerLeuProValLeuAsnSerPheValSerLysArgVal 174
 QY 743 GATGACATCTGCTGCTGCTGGCTTGGAGAGAACTAAGATTTTCACAGACGTATAT 802
 Db 175 AspArgMetLeuGluAlaGlyLeuLeuGluGluValArgGluVal-----PheAsn 191
 QY 803 CAGAGAAATGTTTGGAAATATAGCCAGACTATCAACATGATCTTCCATCAATATGCG 862
 Db 192 ProLys-----AlaAsnTyrSerValGlyIleArgArgAlaIleGly 205
 QY 863 TTCAGGAATTTACAGAGTACCTGATCACTGAGGGAATATGACACTGAGACTGATAC 922
 Db 206 ValProGluLeuHisGluTyrLeuArgAsnGluSerLeuValAspArgAlaThrLysSer 225
 QY 923 CAGCTTCTA-----AAGAAAGACCTGGTCCATTGTC----- 955
 Db 226 LysMetLeuAspValAlaValLysAsnIleLysLysAsnThrGluIleLeuAlaCysArg 245
 QY 956 -----CCCCCTGTCTATGCTTAGAG 976
 Db 246 GlnLeuLysLysIleGlnArgLeuHisLysLysTyrLysMetSerMetHisArgValAsp 265
 QY 977 GTATCTGATGTC-----TCGAAGTGGGAGAGCT 1006
 Db 266 AlaThrGluValPheLeuLysArgAsnValGluGluGlnAspGluAlaThrProLysAsnLeu 285
 QY 1007 GTTCTTGAACCTGCTCTGAAATCGTCAAGTTTCATCCAGGCGCAAGCTACAGCC 1066
 Db 286 ValAlaArgProSerGluArgIleValAspLysPheTyrAsnAsnAsn----- 301
 QY 1067 ACTCAATAAAGATGCCATACATGAAAGCTGAGACAGACAGAAAGTTATACCTGTGTGAC 1126
 Db 302 -----AsnGlnLeuLysAsnAspAspValGluHis-Cys----- 312
 QY 1127 CTCTGTGATGCAATCATCTATGGGATGCCGAATGGCGCACCACTAAAT 1178
 Db 313 -----LeuAlaAlaSerTyrGlyLysSerArgAlaHisAsn 327
 RESULT 13
 ID ABB92405 standard; Protein; 329 AA.
 AC ABB92405;
 XX
 DT 31-MAY-2002 (first entry)

```
DE xx Herbicidally active polypeptide SEQ ID NO 1616.  
KW xx Herbicidal; plant; agriculture; herbicide.  
OS xx Arabidopsis thaliana.  
PN xx W0200210210-A2.  
PD xx  
PE xx 07-FEB-2002.  
PF xx 28-AUG-2001; 2001WO-EP09892.  
PI xx 28-AUG-2001; 2001WO-EP09892.  
PP xx (FARB ) BAYER AG.  
PT xx Tietjen K., Weldler M;  
PR xx WPJ; 2002-269010/31.  
PS xx Identifying plant target proteins for herbicidally active compounds,  
XX comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms -  
YY Claim 5; SEQ ID NO 1616; 261pp + Sequence Listing; English.
```

```
The invention relates to identifying target proteins  
(ABB96790-ABB94016) for herbioidally active compounds, comprising  
aligning and comparing nucleic acid or amino acid sequences from plant  
with nucleic acid or amino acid sequences from non-plant organisms using  
suitable search parameters, where plant sequences having an E-value  
greater by a factor of 3 than the E-value of most similar non-plant  
sequences are selected. The polypeptides or nucleic acids encoding them  
are useful as herbicides. The identified modulators are
```

```
Alignment Scores:  
Pred. No.:      1,06e-26          Length:           329  
Score:         333.50             Matches:            99  
Percent Similarity:   38.94%       Conservative:        56  
Best Local Similarity: 24.87%     Mismatches:         105  
Query Match:       9.33%         Indels:              138  
DB:                23               Gaps:                 9
```

```
US-09-513-151-3 (1-2041) x ABB92405 (1-329)  
QY 83 CTTGTAGTATTCCTCGGGCCACGCGGCACAATCACCCTGCCGCTTAGCAAGC 142  
Db : ::::: ||||| ::||| |::::: ||| ::::: ||  
Dd Valillephevalmetgylalathncllysercilysserarglenualalleaspleuala 55  
QY 143 CAGCGGCTCGGGTGATGCTGATCGCATCGAAGCTCTATGAAGCCCTAGAC 202  
Db : ||| ||||||| ::||| ::||| ::||| ::||| ::|||  
Dd Thrartphgehnlgynluilelleanssearsphysilleglnleuryrlsgyleuasasp 75  
QY 203 ATCATCACCAAAGAATTTCGCCAACAGACAGCAATTCGCCGACCCAAATGACCAAGC 262  
Db : ::::: ::||| ::||| ||| ||||| ::||| ::||| ::||| ::|||  
Dd Valleurtnrnsnlvsvalthrproylselucysrglalyalprhoishstlseuleugly 95  
QY 263 TTGTGATCCTCTTGTGACCAATTAACACAGTGATGACTTCAGAAATAGACCACTGCT 322  
Db : ||| ||| ||| ||| ::||| ::||| ::||| ::||| ::|||  
Dd Valpheaspserrguialaglyasnleuthralathtcnlofryserarfglenallaseargin 115  
QY 323 CTGATTGAAGATAATTGCCCCGAGACAATAATCTATTGTTGGAGAGAACCAATTAT 382  
Db : ||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||  
Dd Alaleesertlysleseralaaasnslsylseupeurofillevallaiaeglylserrasen 135  
QY 383 TAGATTAATCTCTGCGCTCGAAMAAGTTCTGCAATACCAAGCCCAGAGGAGTGGCACT 442  
Db : ||||| ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||  
Dd Tyrtllgegulaalabu----- 140
```

OY	443	GAGAAAGTATGACCCAAAAGTGGAGCTTGAAAAGAGAGATGGCTTGTACTTACAAA	502
Db	140	-----	140
OY	503	CGCCTAAGCCAGGTGGACCCAGAAATGGCTGGCCAGCTGCATTCACATGACAAAGCAAA	562
Db	140	-----	140
OY	563	GTGGCCAGAGCCTTGCAAATTTTGAAGAAACAGAAATCTTCATAGTGAATTTCTGCAT	622
Db	141	-----ValAsnHisSerSer-----	145
OY	623	CGTCAACATACGAGAAAGGTGGTGGCTCCCTTGGAGGCTCTCTGAAGTCTCTAACCT	682
Db	146	-----GlyPheLeuLeuAsnAsnTyrIAspCys	154
OY	683	TGCATCCTTTGGCTTCATGCTGCACAGCCAGGCTTCTGATGAGCGGCTTGATGAAGGGTG	742
Db	155	CysPheIleTrrPValAsrPValSerLeuProValLeuAsnSerPheValSerIysArgVal	174
OY	743	GATGACATCTTGGCTGCTGGCTCTTGGAGAACTAAGAGATTTTTCACAGACGCTATAT	802
Db	175	AspArgMetMetGluAlaGlyLeuLeuGluGluValArgGluVal-----PheAsn	191
OY	803	CAGAGAAATGTTTGGAAATTACCACAGACTATCAACATGGATCTTCCATCAATCATGCG	862
Db	192	ProLys-----AlaAsnTyrSerValGlyIleArgTrgAlaIleCyl	205
OY	863	TTCAAGGAATTTCAAGAGTACTGCATCAGTACAGAGGAAATGCAACACTGAGACSTAGTAAC	922
Db	206	ValProGluLeuHisGlnTyrIleuValArgAsnGlySerLeuValAsrArgAlaThrLysSer	225
OY	923	CACGTTCTA-----AAGAAAGACCTGGTCCATTGTC-----	955
Db	226	LysMetLeuAspValAlaValLysAsnIleLysLysAsnThrGluIleLeuAlaCysArg	245
OY	956	-----CCSCCTGTATGGCTTAGAG	976
Db	246	GlnLeuLysLysIleGlnArgLeuHisLysLysTrpLysMetSerMetHisArgValAsp	265
OY	977	GTAATCGATGTC-----TCAAGTGGAGAGAGTCT	1006
Db	266	AlaThrGluValPheLeuLysArgAsnValGluGluAsrPglValAlaTrpLysLeu	285
OY	1007	GTTCTTGAACCTGCTTGAATTCGTGCAAGTTTATCTCAGGGCCACAGCCTACAGCC	1066
Db	286	ValAlaArgProSerGluArgIleValAsrPlysPheTyrAsnAsnAsn-----	301
OY	1067	ACTGCATTAAGAATGCATCATCAAGAGCGAGAGCAACAAAGAGTAAATTCCTGTGTGAC	1126
Db	302	-----AsnGlnLeuLysAsnAsrAspValGluHisCys-----	312
OY	1127	CTCTGTGATCGAATCATCATTTGGGATCCCGAATGGGACGCCACATAAAT	1178
Db	313	-----LeuAlaIaSerTyrGluGlyIysSerGlySerAlaGlnHisAsn	327
RESULT 14			
ID ABB53936 standard; Protein; 294 AA.			
AC ABB53936;			
XX 16-MAY-2002 (first entry)			
DE Lactococcus lactis protein miaa.			
KM Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.			
OS Lactococcus lactis IL1403.			
XX FR2807446-A1.			
XX			

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139432.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140625.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144371.
PR 19-JUL-1999; 99US-0144371.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

```

PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161922.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.

Alignment Scores:
Pred. No.: 4,38e-26 Length: 290
Score: 327.50 Matches: 99
Percent Similarity: 38.58% Conservative: 53
Best Local Similarity: 25.13% Mismatches: 104
Query Match: 9.16% Indels: 138
DB: 21 Gaps: 9

US-09-513-151-3 (1-2041) x AAC49766 (1-290)
OY 95 CTGGGGCCACGGGCGGCAATCCAGCGCTTGCGAGCTAGCGCGGCTGCGC 154
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MetGlyAlaThrGlySerGlyLysSerArgLeuAlaIleAspLeuAlaThrArgPheGln 20

OY 155 GGTGAGATCGTCAGCGCTGCTCCATGCGAGCTATAGAGCGCTAGACATCACCAC 214
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 21 GlyGluIleIleAsnSerAspLysIleGlnLeuTyrLysGlyLeuAspValLeuThrAsn 40

OY 215 AAGTTTCTGCCAAGAGAGAGATCTGCCGCGCACATGATGATCCTTGCGATCCT 274
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 41 LysValThrProLysGluCysArgGlyValProHisIleuGlyValPheAspSer 60

OY 275 CTGTGACCAATTACACAGTGGTGAGCTTGCAAAATAGACAACCTCTGATTAAGAT 334
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GluAlaGlyAsnLeuThrAlaThrGlnTyrSerArgLeuAlaSerGlnAlaIleSerLys 80

OY 335 ATATTGCCCGGACAGCAAAATCTCTATTGTGTGGAGGAGCAATTTTACATTTGATCT 394
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 81 LeuSerAlaAsnAsnLysLeuProIleValIleGlySerAsnSerTyrIleGluAla 100

OY 395 CTGCTCTGAAAGTTCTTGTCAATACCAACCCAGAGAGATGGCACTGAGAAAGTGATT 454
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 101 Leu----- 101

OY 455 GACCGAAAGTGAGCTTGAAAGAGAGATGCTTGTACTTCAAAAGCCTAAGCCAG 514
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 101 ----- 101

OY 515 GTGGACCCAGAAATGGCTGCGCAAGCTGCATCCACATGACAACGAAAGTGCCAGAGC 574
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 101 ----- 101

OY 575 TTGCAAGTTTGAAGAAACAGAAATCTCTATAGTGAATTTCTCCATCTCAACATACG 634
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 102 -----ValAsnHisSer----- 106

OY 635 GAAGAAGTGCTGTCCTTGAGAGTCTCTGAAGTCTCTAACCCTTGCACTCTTGG 694
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 107 -----GlyPheLeuLeuAsnAsnTyrAspCysPheIleTyr 119

OY 695 CTTGATGCTGACGAGCAGCTTCTAGATGAGCGCTTGATAGAGGCTGATGACATGCTT 754
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 ValAspValSerLeuProValIleuAsnSerPheValSerLysArgValAspArgMetMet 139

OY 755 GCTGCTGGGCTTGTGAGAACTAAGAGATTTTCACAGACGCTATATACAGAAATGTT 814
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 140 GluAlaGlyLeuLeuGluValArgGluVal-----PheAsnProLys----- 154

OY 815 TCGGAAATAGCCAGAGCATCAACATGCTATCTCCAAATGAGCTTCAAGGAATTT 874
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 155 -----AlaAsnTyrSerValGlyIleArgArgAlaIleGlyValProGluLeu 170

```

```

OY 875 CACGAGTACCTGATCACTAGGAGGAAATGCACACTGAGACTAGTACCACTTCTA--- 931
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 171 HisGluTyrLeuArgAsnGluSerLeuValAspArgAlaThrLysSerLysMetLeuAsp 190

OY 932 -----AAGAAAGACCTGGTCCCATTTGTC----- 955
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 191 ValAlaValLysAsnIleLysLysAsnThrGluIleLeuAlaCysArgGlnLeuLysLys 210

OY 956 -----CCCCCTGCTATGCGCTTAGAGGTAATCTGATGTC 988
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 211 IleGlnArgLeuHisLysLysTyrPlyMetSerMetHisArgValAspAlaThrGluVal 230

OY 989 -----TCGAAGTGGAGAGAGCTGTCTTGAACT 1018
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 231 PheLeuLysArgAsnValGluGluGlnAspGluAlaThrGluAsnLeuValAlaArgPro 250

OY 1019 GCTCTGAAATCGTCAAAAGTTTCATCCAGGCGCACAAAGCCTACAGCCACTCCAAATAAG 1078
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 251 SerGluArgIleValAspLysPheTyrAsnAsn----- 262

OY 1079 ATGCCATACATGAAGCTGAGAACAGAGAGATTATCAGCTGTGTGACCTGTGTGATCGA 1138
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 263 -----AsnGlnLeuLysAsnAspAspValGluHis-Cys-----LeuAl 275

OY 1139 ATCATCATTTGGGATCGCGAATGGGCGAGCCGACATTAAT 1178
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 275 AlaIleSerTyrGlyGlyLysSerArgArgAlaHisAsn 288

```

Search completed: April 21, 2003, 18:48:16
 Job time : 80.8325 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 19:14:35 : Search time 70 Seconds

(without alignments) 9162.495 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 661

Sequence: 1 CTGCATAGATGCGCTCCG.....TTTACACAGAAAAA 2041

Scoring table:

OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 714767 seqs, 157122656 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1386358

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlh
-O/cgn2_1/USPTO_spool/US09513151/runat_15042003_141248_27291/app_query.fasta_1.2183
-DB=Pending_Patents_AA_New -OFMT=fasta -SUFFIX=olin2p.rapn -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=pio -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09513151 -ECGN_1.1_49 -runat_15042003_141248_27291
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Pending Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	254	38.4	411	6	US-10-380-731-563
2	203	30.7	475	7	US-60-453-135-7968
3	203	30.7	475	7	US-60-453-050-7968
4	109	16.5	221	7	US-60-453-135-7967
5	109	16.5	221	7	US-60-453-050-7967
6	63	9.5	85	7	US-60-453-135-7969
7	63	9.5	85	7	US-60-453-050-7969
8	48	7.3	57	6	US-10-203-138A-11847
9	9	1.4	254	6	US-10-366-683-25445
10	9	1.4	365	6	US-10-369-493-8141
11	9	1.4	387	6	US-10-366-683-16934

12	9	1.4	713	6	US-10-156-761-13800	Sequence 13800, A
13	9	1.2	124	6	US-10-282-122A-48736	Sequence 48736, A
14	8	1.2	136	5	US-09-864-408A-124	Sequence 124, App
15	8	1.2	200	6	US-10-366-683-18739	Sequence 18739, A
16	8	1.2	212	1	PCT-US02-32727-26700	Sequence 26700, A
17	8	1.2	212	5	US-09-978-825-26700	Sequence 26700, A
18	8	1.2	212	6	US-10-057-498-26700	Sequence 26700, A
19	8	1.2	216	6	US-10-282-122A-55271	Sequence 55271, A
20	8	1.2	216	6	US-10-156-761-9951	Sequence 9951, A
21	8	1.2	219	1	PCT-US02-32727-8832	Sequence 8832, App
22	8	1.2	219	5	US-09-978-825-5852	Sequence 5852, App
23	8	1.2	219	6	US-10-057-498-5852	Sequence 5852, App
24	8	1.2	248	5	US-09-134-000C-4333	Sequence 4333, App
25	8	1.2	248	5	US-09-134-000C-4333	Sequence 4333, App
26	8	1.2	249	6	US-10-366-683-30556	Sequence 30556, A
27	8	1.2	257	1	PCT-US02-36123-2920	Sequence 2920, App
28	8	1.2	269	1	PCT-US02-32727-26977	Sequence 26977, A
29	8	1.2	269	5	US-09-978-825-26977	Sequence 26977, A
30	8	1.2	269	6	US-10-057-498-26977	Sequence 26977, A
31	8	1.2	288	1	PCT-US02-24310-77	Sequence 77, App1
32	8	1.2	295	5	US-09-724-676-76506	Sequence 76506, A
33	8	1.2	295	5	US-09-724-676-76506	Sequence 76506, A
34	8	1.2	305	6	US-10-366-683-17702	Sequence 17702, A
35	8	1.2	324	7	US-60-453-135-9949	Sequence 9949, App
36	8	1.2	324	7	US-60-453-050-9949	Sequence 9949, App
37	8	1.2	357	6	US-10-282-122A-46768	Sequence 46768, A
38	8	1.2	370	6	US-10-366-683-18438	Sequence 18438, A
39	8	1.2	370	6	US-10-366-683-18438	Sequence 18438, A
40	8	1.2	378	5	US-09-989-733-156	Sequence 156, App
41	8	1.2	378	5	US-09-992-643-156	Sequence 156, App
42	8	1.2	378	6	US-10-131-813A-420	Sequence 420, App
43	8	1.2	378	6	US-10-131-819A-420	Sequence 420, App
44	8	1.2	378	6	US-10-131-823A-420	Sequence 420, App
45	8	1.2	378	6	US-10-131-824A-420	Sequence 420, App

ALIGNMENTS

RESULT 1
US-10-380-731-563
Sequence 563, Application US/10380731
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
FILE REFERENCE: 21272-114
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
CURRENT APPLICATION NUMBER: US/10/380,731
CURRENT FILING DATE: 2003-03-14
PRIOR APPLICATION NUMBER: 09/659,671
PRIOR FILING DATE: 2000-09-11
NUMBER OF SEQ ID NOS: 888
SOFTWARE: Custom
SEQ ID NO 563
LENGTH: 411
TYPE: PRT
ORGANISM: Homo sapiens
US-10-380-731-563

Alignment Scores:

Pred. No.: 4.81e-254
Score: 254.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 38.43%
DB: 6
Length: 411
Matches: 254
Conservative: 0
MisMatches: 0
Indels: 0
Gaps: 0

US-09-513-151-3 (1-2041) x US-10-380-731-563 (1-411)

OY 179 ATGCAGGTATGAGAGCCTAGACATCATCACAAGGTTCTGCCAGACAGAGA 238
|||||
Db 1 MergInValTfclncllyleuaspIleIethAsnlylValserIadInclngInag 20
OY 239 ATCTGCGGCGCACCATGATCAGCTTGTGATCTCTTGTGACCAATTACAGTGCTG 298
|||||

Db 21 IleCysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrThrValVal 40
OY 299 GACSTCAGAAATAGAGCAACTGCTGTGATGAGATATATATGCCCCGAGACAAATTCCT 358
Db 41 ASPPheAAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIlePro 60
OY 359 ATTGTGTGGAGCAACCAATATATACATTCATGCTGCTGCTGCAAGATTCCTTCAT 418
Db 61 IleValValGlyGlyThrAsnTyrTyrIleGluSerLeuLeuTyrLysValLeuValAsn 80
OY 419 ACCAAGCCCCAGAGATGGGCACTGAGAAAGTATGACCGAAAGTGGAGCTGAAAG 478
Db 81 ThrLysProGlnGluMetGlyThrCyluLysValIleAspArgLysValGluLeuGluLys 100
OY 479 GAGATGGCTGTGACTTCACAAACGCTTAAGCCGAGCTGGACCCGAAATGGCTCCAG 538
Db 101 GlnuSpGlyLeuValLeuHisLysArgLeuSerGlnValAspProGluMetAlaLys 120
OY 539 CTGCATCCACATGACAAACGCAAGTGGCCAGAGACTTGAAGTTTGAAGAAGACAGA 598
Db 121 LeuHisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGluGluThrGly 140
OY 599 ATCTCTCATGTAATTTCTCCATCGTCACATACGGAAGAAGGTGTGCTCCCTTGA 658
Db 141 IleSerHisSerGluPheLeuHisArgGlnHisThrGluGluGlyGlyProLeuGly 160
OY 659 GGTCCCTCGAAGTCTCTCAACCTTGCATCTTGGCTTCATGCTGACAGCACTCTTA 718
Db 161 GlnProLeuLysPheSerAsnProCysIleLeuTyrPheHisAlaAspGlnAlaValLeu 180
OY 719 GATGAGCCCTTGATTAAGAGGTGATGACATGCTTGGCTGGGCTCTTGAGAGACTA 778
Db 181 AspGluArgLeuAspLysArgValAlaAspPheLeuAlaIleGlyLeuLeuGluGluLeu 200
OY 779 AGAGATTTTCACAGACGCTAATATATCAGAAATGTTTGGAAATATAGCCAGACTATCA 838
Db 201 ArgAspPheHisArgArgTyrAsnGlnLysAsnValSerGluAsnSerGlnAspTyrGln 220
OY 839 CATGATATCTCCAAATCAATGCGCTTCAGAAATTCAGAGATACCTATCACTAGAGGA 898
Db 221 HisGlyIlePheGlnSerIleGlyPheLysGluPheHisGluTyrLeuIleThrGluGly 240
OY 899 AATGACACACTGAGACACTAGTACCAGCTTCTAAAGAAGA 940
Db 241 LysCysThrLeuGluThrSerAsnGlnLeuLeuLysGly 254

RESULT 2
US-60-453-135-7968
: Sequence 7968, Application US/60453135
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: IAKOUBOVA, Olga
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C0001456
: CURRENT APPLICATION NUMBER: US/60/453,135
: CURRENT FILING DATE: 2003-03-10
: NUMBER OF SEQ ID NOS: 82762
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 7968
: LENGTH: 475
: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-453-135-7968

Alignment Scores:
Pred. No.: 3,87e-201
Score: 203.00
Percent Similarity: 100.008
Best Local Similarity: 100.008
Query Match: 30.71%
DB: 7
Length: 475
Matches: 203
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-513-151-3 (1-2041) x US-60-453-135-7968 (1-475)
OY 332 CATATATTTGCCCCGAGACAAATTCCTATGCTTGGGAGAGCAACCAATATATACATGAA 391
Db 116 ASPillePheAlaArgAspLysIleProIleValIleGlyGlyThrAsnTyrTyrIleGlu 135
OY 392 TCTGCTGCTGGAAAGTCTTGTCAATACCAAGCCCGAGATGGGCACTGAGAAAGT 451
Db 136 SerLeuLeuTyrLysValLeuValAsnThrLysProGlnGluMetCyluThrGluLysVal 155
OY 452 ATTGACGAAAAGTGGAGCTTGAAGAAGAGATGGCTTGTATTCACAAACGCTTAAGC 511
Db 156 IleAspArgLysValGluLeuGluLysGluAspArgLysValLeuHisLysArgLeuSer 175
OY 512 CAGGTGGACCCGAAATATGCTGCTCCAAAGTGCATGCACATGACAAACGCAAGTGGCCAG 571
Db 176 GlnValAspProGluMetAlaIleValLysLeuHisProHisAspLysArgLysValAlaArg 195
OY 572 AGCTTGCAAGTTTGAAGAAGACAGATCTCTCATGTAATTTCTGCATGCTGACAT 631
Db 196 SerLeuGlnValPheGluGluThrGlyIleSerHisSerGluPheLeuHisArgGlnHis 215
OY 632 ACGGAAGAAGTGTGTGCTCCCTTGGAGGTCTCTGAAGTTCCTTAACCTTGGACTCTT 691
Db 216 ThrGluGluGlyGlyProLeuGlyGlyProLeuLysPheSerAsnProCysIleLeu 235
OY 692 TGCTTCACTGCTGACCCAGGCACTTACATGAGCCCTTGATTAAGAGGTGATACATG 751
Db 236 ThrLeuHisAlaAspGlnAlaValLeuAspArgLysValAlaAspPhe 255
OY 752 CTGCTGCTGGGCTTGGAGAACTAAGATTTTCACAGACGCTAATATCAGAAAGAT 811
Db 256 LeuAlaIleGlyLeuLeuGluGluLeuLeuArgAspPheHisAlaArgTyrAsnGlnLysAsn 275
OY 812 GTTTCGAAAATAGCCAGACTATACATGATGATCTTCAATCAATTGGCTTCAAGGAA 871
Db 276 ValSerGluAsnSerGlnAspTyrGlnHisGlyIlePheGlnSerIleGlyPheLysGlu 295
OY 872 TTTTCAGAGTACCTGATACCTAGAGGAAATATGACACTGAGACTAGTACCAGCTTGA 931
Db 296 PheHisGluTyrLeuIleThrGluGluLysCysThrLeuGluThrSerAsnGlnLeuLeu 315
OY 932 AAGAAGA 940
Db 316 LysLysGly 318

RESULT 3
US-60-453-050-7968
: Sequence 7968, Application US/60453050
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: LUKE, May
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C0001457
: CURRENT APPLICATION NUMBER: US/60/453,050
: CURRENT FILING DATE: 2003-03-10
: NUMBER OF SEQ ID NOS: 82762
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 7968
: LENGTH: 475
: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-453-050-7968

Alignment Scores:
Pred. No.: 3,87e-201
Score: 203.00
Percent Similarity: 100.008
Best Local Similarity: 100.008
Query Match: 30.71%
DB: 7
Length: 475
Matches: 203
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-513-151-3 (1-2041) x US-60-453-050-7967 (1-475)

```

OY 332 GATATATTGCCCCGAGCAAAATTCCTATTGTTGGAGAGCAACATATTACATTGAA 391
    |||||||
Db 116 AsplIeHehIaIarGAsPlYsIleProIleValIaGIgYlThrAsTyrTrIleGlu 135
OY 392 TCTGTGCTGTGGAAGTTCTTGTCAATACCAAGCCCCGAGAGGGGACACGAGAAAGTG 451
    |||||||
Db 136 SerLeuLeuTrPlYsValLeuValAsnThrLysProGInGluMetGlyThGluYsVal 155
OY 452 ATTACCGCAAAAGTGAGACTTGAAAGAGAGATGGCTTGTACTTCAACAAAGCCCTAAGC 511
    |||||||
Db 156 IleAspArGlySValGleuGluYsGluSlnsPrgIYLeuValIleuHISlyArGLeuSer 175
OY 512 CAGGTGAGACCCAGAAATGGCTGCCAAGCTGCATCCACATGACAAACGCAAACTGGCCAGG 571
    |||||||
Db 176 GluValAsPProGluMeValaIaIaLysLeuHISProHISAsPlYsArGlySValIlaArG 195
OY 572 AGCTTGCAAGTTTGAAGAAACAGAGATCTCTCATAGTGAATTTCTCCATGCTCAACAT 631
    |||||||
Db 196 SerLeuGlnValaPheGlnGluThrGlyIleSerHISSerGluPheLeuHISArGlnHIS 215
OY 632 AGCGAAGAGGTGGTGGTCCCTTGAGAGCTCTGAAAGTTCTTAACCTTGATCCTT 691
    |||||||
Db 216 ThrGlnGluGlyGlyGlyProLeuGlnGlyGlyProLeuLysPheSerAsnProCysIleLeu 235
OY 692 TGAGCTTCATGCTGACAGGACAGTTCTAGATGAGCGCTTGATAGAGGCTGATGACATG 751
    |||||||
Db 236 ThrLeuHISAlaAspGlnIlaValaIleuAspGluArGLeuAsPlYsArGlyValaAspArMet 255
OY 752 CTGTGCTGTGGCTTGTGAGAGCACTAAGAGATTTTCACAGAGGCTATATCGAAGAAAT 811
    |||||||
Db 256 LeuAlaAlaGlyLeuLeuGlnGluLeuValaGAsPheHISArGlyTrAsnGlnLysAsn 275
OY 812 GTTTCGGAATATGCCAGAGCATATCAACATGATCTTCCAAATCAATTTGGCTTCAGAGAA 871
    |||||||
Db 276 ValSerGluAsnSerGlnAspTyrGlnHISGlyIlePheGlnSerIleGlyPheLysGlu 295
OY 872 TTTCACAGTACTGATCACTGAGAGGAAATGCACTGAGAGCTAGTAAACGACTTCTTA 931
    |||||||
Db 296 PheHISGluTrIleuIleThrGlnGlyLysCysThrLeuGluThrSerAsnGlnLeuLeu 315
OY 932 AAGAAAGCA 940
    |||||||
Db 316 LysLysGly 318

```

RESULT 4

US-60-453-135-7967
Sequence 7967, Application US/60453135

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: IAKOUBOVA, Olga

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001456

CURRENT APPLICATION NUMBER: US/60/453,135

CURRENT FILING DATE: 2003-03-10

NUMBER OF SEQ ID NOS: 82762

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7967

LENGTH: 221

TYPE: PRT

ORGANISM: Homo sapiens

US-60-453-135-7967

Alignment Scores:

Pred. No.: 1.45e-103 Length: 221
Score: 109.00 Matches: 109
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.49% Indels: 0
DB: 7 Gaps: 0

US-09-513-151-3 (1-2041) x US-60-453-135-7967 (1-221)

```

OY 941 CCTGTCCCATTTGCCCCCTGTATAGGCTTAGAGGATCTGATGCTGCAAGTGGAG 1000
    |||||||
Db 91 ProGlyProIleValaProProValaIYrGlyLeuGluValSerAspValSerLysTrpGlu 110
OY 1001 GAGTCTGTCTTGAACCTGCTCTTGAATCTGCAAACTTTCATCCAGGCCACAAAGCT 1060
    |||||||
Db 111 GluSerValaLeuGluProAlaLeuGluIleValaGlnSerPheIleGlnGlyHISLysPro 130
OY 1061 ACAGCCACTCCCAATTAAGATGCCATACATGAGCTGAGAACAGAGAGAAATTAACCTG 1120
    |||||||
Db 131 ThrAlaThrProIleLysMetProGlyAsnGluAlaGlnAsnLysArgSerTrpHISLeu 150
OY 1121 TGTGACCTGTGTATGCAATATCTATGCGGAGATCGCGAATGGGACCCACATAAATCC 1180
    |||||||
Db 151 CysAspLeuCysAspArGlyIleIleIleGlyAspArGlyTrpAlaIaHISLysSer 170
OY 1181 AAATCCCACTTGACCACTGAGAAAGAAAGAAAGATTGACTCAGATGCTGCAACAC 1240
    |||||||
Db 171 LysSerHISLeuAsnGlnLeuLysLysArGArGArGLeuAspSerAspAlaValaSnThr 190
OY 1241 ATGAAAGTCAAGTGTTCGCCAGAC 1267
    |||||||
Db 191 IleGluSerGlnSerValSerProAsp 199

```

RESULT 5

US-60-453-050-7967
Sequence 7967, Application US/60453050

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: LUKE, May

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001457

CURRENT APPLICATION NUMBER: US/60/453,050

CURRENT FILING DATE: 2003-03-10

NUMBER OF SEQ ID NOS: 82762

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7967

LENGTH: 221

TYPE: PRT

ORGANISM: Homo sapiens

US-60-453-050-7967

Alignment Scores:

Pred. No.: 1.45e-103 Length: 221
Score: 109.00 Matches: 109
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.49% Indels: 0
DB: 7 Gaps: 0

US-09-513-151-3 (1-2041) x US-60-453-050-7967 (1-221)

```

OY 941 CCTGTCCCATTTGCCCCCTGTATAGGCTTAGAGGATCTGATGCTGCAAGTGGAG 1000
    |||||||
Db 91 ProGlyProIleValaProProValaIYrGlyLeuGluValSerAspValSerLysTrpGlu 110
OY 1001 GAGTCTGTCTTGAACCTGCTCTTGAATCTGCAAACTTTCATCCAGGCCACAAAGCT 1060
    |||||||
Db 111 GluSerValaLeuGluProAlaLeuGluIleValaGlnSerPheIleGlnGlyHISLysPro 130
OY 1061 ACAGCCACTCCCAATTAAGATGCCATACATGAGCTGAGAACAGAGAGAAATTAACCTG 1120
    |||||||
Db 131 ThrAlaThrProIleLysMetProGlyAsnGluAlaGlnAsnLysArgSerTrpHISLeu 150
OY 1121 TGTGACCTGTGTATGCAATATCTATGCGGAGATCGCGAATGGGACCCACATAAATCC 1180
    |||||||
Db 151 CysAspLeuCysAspArGlyIleIleIleGlyAspArGlyTrpAlaIaHISLysSer 170
OY 1181 AAATCCCACTTGACCACTGAGAAAGAAAGAAAGATTGACTCAGATGCTGCAACAC 1240
    |||||||

```

Db 171 LysSerHisLeuAsnGlnLeuLysArgArgArgLeuAspSerAspAlaValAsnThr 190
 QY 1241 ATGAAAGTCAGACTGTTTCCCGCAGC 1267
 Db 191 lIeGlSerGlnSerValSerProAsp 199

RESULT 6

US-60-453-135-7969
 ; Sequence 7969, Application US/60453135
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; APPLICANT: IAKOUBOVA, Olga
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: C1001456 MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
 ; CURRENT APPLICATION NUMBER: US/60/453,135
 ; NUMBER OF SEQ ID NOS: 82762
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 7969
 ; LENGTH: 85
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-60-453-135-7969

Alignment Scores:

Pred. No.: 8,76e-56 Length: 85
 Score: 63.00 Matches: 63
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 9.53% Indels: 0
 DB: 7 Gaps: 0

US-09-513-151-3 (1-2041) x US-60-453-135-7969 (1-85)

QY 1079 ATGCCATACATGACCTGAGAACAGAGAGTATCACCCTGCTGTGATCGA 1138
 Db 1 MetProTyrAsnGlnAlaGlnAsnLysArgSerTyrHisLeuCysAspArg 20
 QY 1139 ATCATCATGTCGATCGCGAATGGCGACGCGACATTAATCCCAATCCCACTTGAACCA 1198
 Db 21 lIeIleIleGlyAspArgGlnTyrPalaAlaHisIleLysSerLysSerHisLeuAsnGln 40
 QY 1199 CTGAAGAAAGAAGAAGATTGACATGCTGTCAACACCATGAAGTGAAGTGT 1258
 Db 41 LeuLysLysArgArgArgLeuAspSerAspAlaValAsnThrIleGlSerGlnSerVal 60
 QY 1259 TCCCGCAGC 1267
 Db 61 SerProAsp 63

RESULT 7

US-60-453-050-7969
 ; Sequence 7969, Application US/60453050
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; APPLICANT: LUKE, May
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: C1001457 STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; CURRENT APPLICATION NUMBER: US/60/453,050
 ; NUMBER OF SEQ ID NOS: 82762
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 7969
 ; LENGTH: 85
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-60-453-050-7969

Alignment Scores:

Pred. No.: 8,76e-56 Length: 85
 Score: 63.00 Matches: 63

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 9.53% Indels: 0
 DB: 7 Gaps: 0

US-09-513-151-3 (1-2041) x US-60-453-050-7969 (1-85)

QY 1079 ATGCCATACATGACCTGAGAACAGAGAGTATCACCCTGCTGTGATCGA 1138
 Db 1 MetProTyrAsnGlnAlaGlnAsnLysArgSerTyrHisLeuCysAspArg 20
 QY 1139 ATCATCATGTCGATCGCGAATGGCGACGCGACATTAATCCCAATCCCACTTGAACCA 1198
 Db 21 lIeIleIleGlyAspArgGlnTyrPalaAlaHisIleLysSerLysSerHisLeuAsnGln 40
 QY 1199 CTGAAGAAAGAAGAAGATTGACATGCTGTCAACACCATGAAGTGAAGTGT 1258
 Db 41 LeuLysLysArgArgArgLeuAspSerAspAlaValAsnThrIleGlSerGlnSerVal 60
 QY 1259 TCCCGCAGC 1267
 Db 61 SerProAsp 63

RESULT 8

US-10-203-138A-11847
 ; Sequence 11847, Application US/10203138A
 ; GENERAL INFORMATION:
 ; APPLICANT: Molecular Dynamics, Inc.
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
 ; FILE REFERENCE: PB 0004 WO 8
 ; CURRENT APPLICATION NUMBER: US/10/203,138A
 ; PRIOR FILING DATE: 2002-08-02
 ; PRIOR FILING DATE: 04 February 2000 (04.02.00)
 ; PRIOR FILING DATE: 26 May 2000 (26.05.00)
 ; PRIOR FILING DATE: 03 August 2000 (03.08.00)
 ; PRIOR FILING DATE: 03 October 2000 (03.10.00)
 ; PRIOR FILING DATE: 27 September 2000 (27.09.00)
 ; PRIOR FILING DATE: 21 September 2000 (21.09.00)
 ; PRIOR FILING DATE: 30 June 2000 (30.06.00)
 ; NUMBER OF SEQ ID NOS: 15438
 ; SOFTWARE: Molecular Dynamics Sequence Listing Engine
 ; SEQ ID NO 11847
 ; LENGTH: 57
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; OTHER INFORMATION: MAP TO AL033527.25
 ; FEATURE: EXPRESSED IN BT474, SIGNAL = 1.1
 ; OTHER INFORMATION: EST_HUMAN HIT: BE242161.1, EVALUE 3.00e-21
 ; OTHER INFORMATION: SWISSPROT HIT: Q9KAC3, EVALUE 2.00e-04
 ; US-10-203-138A-11847

Alignment Scores:

Pred. No.: 3.38e-40 Length: 57
 Score: 48.00 Matches: 48
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.26% Indels: 0

DB: 6 Gaps: 0

US-09-513-151-3 (1-2041) x US-10-369-493-8141 (1-57)

QY 185 GCTATGAAGCCCTAGACATCATCACCAAGAGTTTCTGCCAGACAGATCTGC 244
|||||
DB 1 VAITYRGLUGLYLeuAspIleIleThrAsnLysValSerIleGlnIleGlyCys 20

QY 245 CGGACACATGATCAGCTTGTGGATCTTGTGACCAATTACACAGTGGGACTTC 304
|||||
DB 21 ATGHSISMelIleSerPheValAspProLeuValThrValAspPhe 40

QY 305 AGAATAGACAACTGCTCTGATT 328
|||||
DB 41 ArgAsnArgAlaThrAlaLeuIle 48

RESULT 9

US-10-366-683-25445

Sequence 25445, Application US/10366683

GENERAL INFORMATION:

APPLICANT: Rubenfield, Marc J.

APPLICANT: Nolling, Jork

APPLICANT: Deloughery, Craig

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: PATH03-04

CURRENT APPLICATION NUMBER: US/10/366,683

CURRENT FILING DATE: 2003-02-13

PRIOR APPLICATION NUMBER: 09/252,991

PRIOR FILING DATE: 1999-02-18

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 25445

LENGTH: 254

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-10-366-683-25445

Alignment Scores:

Pred. No.:	8.05	Length:	254
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.36%	Indels:	0
DB:	6	Gaps:	0

US-09-513-151-3 (1-2041) x US-10-366-683-25445 (1-254)

QY 130 GTTCAGCTAGCGCAGCGCTCGCGG 156
|||||

DB 47 ValAlaAlaArgProAlaAlaArgArg 55

RESULT 10

US-10-369-493-8141

Sequence 8141, Application US/10369493

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 8141

LENGTH: 365

TYPE: PRT

ORGANISM: Thermobifida fusca

US-10-369-493-8141

Alignment Scores:

Pred. No.:	7.67	Length:	365
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.36%	Indels:	0
DB:	6	Gaps:	0

US-09-513-151-3 (1-2041) x US-10-369-493-8141 (1-365)

QY 14 GCGTCGTCGCGCTGCAGACAGTT 40
|||||

DB 85 AlaSerValAlaAlaAlaArgAlaVal 93

RESULT 11

US-10-366-683-16934

Sequence 16934, Application US/10366683

GENERAL INFORMATION:

APPLICANT: Rubenfield, Marc J.

APPLICANT: Nolling, Jork

APPLICANT: Deloughery, Craig

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOM

FILE REFERENCE: PATH03-04

CURRENT APPLICATION NUMBER: US/10/366,683

CURRENT FILING DATE: 2003-02-13

PRIOR APPLICATION NUMBER: 09/252,991

PRIOR FILING DATE: 1999-02-18

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 16934

LENGTH: 387

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-10-366-683-16934

Alignment Scores:

Pred. No.:	7.61	Length:	387
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.36%	Indels:	0
DB:	6	Gaps:	0

US-09-513-151-3 (1-2041) x US-10-366-683-16934 (1-387)

QY 98 GGGCCAGCGGCGCAGCAATCCAG 124
|||||

DB 135 GlyAlaThrGlyThrGlyLysSerThr 143

RESULT 12

US-10-156-761-13800

Sequence 13800, Application US/10156761

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 13800


```

; LENGTH: 713
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13800

Alignment Scores:
Pred. No.: 7.01 Length: 713
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 6 Gaps: 0

US-09-513-151-3 (1-2041) x US-10-156-761-13800 (1-713)
OY 95 CTCGGGGCCAGCGGCGCAATCC 121
      |||||
Db 47 LeuGlyAlaThrGlyThrGlyLysSer 55

RESULT 13
US-10-282-122A-48736
; Sequence 48736, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Travick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48736
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Bacteroides fragilis
US-10-282-122A-48736

Alignment Scores:
Pred. No.: 96.6 Length: 124
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```

```

Query Match: 1.21% Indels: 0
DB: 6 Gaps: 0

US-09-513-151-3 (1-2041) x US-10-282-122A-48736 (1-124)
OY 1468 TCCTTGATGCTGCTTTAAAGTCT 1491
      |||||
Db 57 SerPheAspValValLeuLysSer 64

RESULT 14
US-09-864-408A-124
; Sequence 124, Application US/09864408A
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: Novel Human Polynucleotides and Polypeptides Encoded There
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 124
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-124

Alignment Scores:
Pred. No.: 95.4 Length: 136
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 5 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-864-408A-124 (1-136)
OY 113 GGCNAATCCAGCGTGGCGTTCGAC 136
      |||||
Db 14 GlyLysSerThrLeuAlaLeuGln 21

RESULT 15
US-10-366-683-18739
; Sequence 18739, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenstein, Marc J.
; APPLICANT: Nolling, Jock
; APPLICANT: Deloughery, Craig
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMON
; FILE REFERENCE: PATH03-04
; CURRENT APPLICATION NUMBER: US/10/366,683
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18739
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-18739

Alignment Scores:
Pred. No.: 90.6 Length: 200
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 6 Gaps: 0
```

US-09-513-151-3 (1-2041) x US-10-366-683-18739 (1-200)

OY 14 GCGTCGTGCGGCTGCACGACGA 37
|||||
DB 170 AlaservAlaAlaAlaAlaAla 177

Search completed: April 21, 2003, 19:23:31
Job time : 75 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 21, 2003, 19:15:35 ; Search time 274 Seconds

(without alignments)
9605.112 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 661

Sequence: 1 CTGCCATAGATGCGCTCCG.....TTTACAGAGAAAAA 2041

Scoring table:

OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 456914 seqs, 644733110 residues

Word size: 1

Total number of hits satisfying chosen parameters: 8787898

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=x1h
-O=/cgn2_1/USPTO_spool/US09513151/runat_15042003_141248_27274/app_query.fasta.1.2183
-DB=Pending_Patents_AA_Main -QFMT=fastan -SUFFIX=olin2p.ram -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09513151@cgn2_1.1.351@runat_15042003_141248_27274
-NCPU=6 -ICPU=3 -NO_XLPEXT -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database:

Pending_Patents_AA_Main:
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	310	46.9	467	1	PCT-US02-07826-153
2	310	46.9	467	24	US-10-097-340-153
3	274	41.5	465	1	PCT-US02-07826-151
4	274	41.5	465	24	US-10-097-340-151
5	123	18.6	222	1	PCT-US01-01229-1394
6	123	18.6	222	1	PCT-US01-01349-586
7	123	18.6	222	1	PCT-US01-01349-784
8	123	18.6	222	21	US-09-764-853-586
9	123	18.6	222	21	US-09-764-853-784
10	123	18.6	222	21	US-09-764-902-1394
11	123	18.6	222	24	US-10-072-326-586
12	123	18.6	222	24	US-10-072-326-784
13	109	16.5	143	21	US-09-757-028-2199
14	109	16.5	143	26	US-10-222-911-2199
15	48	7.3	57	1	PCT-US01-00663-27779
16	48	7.3	57	22	US-09-864-761-42725
17	48	7.3	57	25	US-10-182-993-26942
18	48	7.3	57	25	US-10-182-997-19892
19	48	7.3	57	26	US-10-203-134-27526
20	48	7.3	57	26	US-10-203-136-27530
21	48	7.3	57	26	US-10-203-137-27779
22	48	7.3	57	26	US-10-203-138-11847
23	48	7.3	57	26	US-10-203-139-26690
24	48	7.3	57	27	US-06-236-359-16268
25	16	2.4	252	16	US-09-270-767-44570
26	10	1.5	4132	23	US-09-980-217-21
27	9	1.4	94	1	PCT-US00-26548-5731
28	9	1.4	143	24	US-10-106-698-5741
29	9	1.4	254	16	US-10-115-127-13676
30	9	1.4	281	21	US-09-252-991A-25445
31	9	1.4	298	25	US-09-793-769-46
32	9	1.4	298	25	US-10-135-881-10491
33	9	1.4	298	26	US-10-219-999-53435
34	9	1.4	365	27	US-06-360-039-8141
35	9	1.4	379	17	US-09-328-352-5804
36	9	1.4	382	21	US-09-791-537-10799
37	9	1.4	386	16	US-09-248-766-17815
38	9	1.4	386	27	US-06-096-409-17815
39	9	1.4	427	16	US-09-252-991A-16934
40	9	1.4	427	21	US-09-791-537-76494
41	9	1.4	428	21	US-09-791-537-46482
42	9	1.4	512	21	US-09-716-865-14
43	9	1.4	698	21	US-09-791-537-83891
44	8	1.2	20	27	US-06-314-359-97
45	8	1.2	39	21	US-09-708-427-72688

ALIGNMENTS

RESULT 1
PCT-US02-07826-153
Sequence 153, Application PC/TUS0207826
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc. et al.
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification, Assessment, Prevention, and Therapy of Ovarian Cancer
FILE REFERENCE: MRI-030PC
CURRENT APPLICATION NUMBER: PCT/US02/07826
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/276,026


```

Db 215 GlyProLeuLysPheSerAsnProCysIleLeuTrpLeuHisAlaAspGlnAlaValLeu 234
QY 719 GATGAGCGCTTGCGATTAAGAGCGTGCATGACATGCTTGCCTGGCGCTTGAGAGACTA 778
Db 235 AspGluArgLeuAspLysArgValAspAspMetLeuAlaIleGlyLeuGluGluLeu 254
QY 779 AGGATTTTCACAGACGCTTAATCAAGAAGATGTTTGGGAAATAGCCAGACTATCA 838
Db 255 ArgAspPheHisArgArgTyrAsnGlnLysAsnValSerGlnAsnSerGlnAspTyrGln 274
QY 839 CATGTAATCTCCAAATTCATTTGGCTTCAAGAAATTTTCAGAGTACTGATCACTAGAGGA 898
Db 275 HisGlyIlePheGlnSerIleGlyPheLysGluPheHisGluTyrLeuIleThrGlnGly 294
QY 899 AAATGCACACTGAGAGACTAGTAACACAGCTTCTTAAGAAGA 940
Db 295 LysCysThrLeuGluThrSerAsnGlnLeuLeuLysGly 308

```

RESULT 4

```

US-10-097-340-151
; Sequence 151, Application US/10097340
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins for The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-151

```

Alignment Scores:

```

Pred. No.: 4,1e-262 Length: 465
Score: 274.00 Matches: 274
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.45% Indels: 0
Db: 24 Gaps: 0

```

```

US-09-513-151-3 (1-2041) x US-10-097-340-151 (1-465)
QY 119 TCCAGCGCTGGCTTGCAGCTAGAGCCAGCGCTGCGCGTGCATGCTCAGCGCTAGCTCC 178
Db 35 SerThrLeuAlaLeuGlnLeuGlnArgGlnArgGlnGlyGlnIleValSerAlaAspSer 54
QY 179 ATCCAGCTATGAGAGCGCTAGACATCATCACCACAAAGCTTTGCCCAAGAGCAGAGA 238
Db 55 MetGlnValTyrGlnIleLysAspIleIleThrAsnLysValSerAlaGlnGlnArg 74
QY, 239 ATCTCCGCGCACACATGATCAGCTTTGGATCCCTCTGTACCAATTAACACAGTGTG 298
Db 75 IleCysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrValVal 94
QY 299 GACTTCAGAAATAGACAACATGCTGATTAAGATATATTATTTGCCCGAGACAAATTCCT 358
Db 95 AspPheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIlePro 114
QY 359 ATTGTTTGGGAGAGAACCAATTATTAATGTAATCTGCTGGAAGTTCTTGTCAAT 418
Db 115 IleValValGlyGlyThrAsnTyrTyrIleGlnSerLeuLeuTrpLysValLeuValAsn 134
QY 419 ACCAAGCCCGAGAGATGGCGCTGAGAAAGATGATGACCCGAAAGTGGACCTGAAAG 478
Db 135 ThrLysProGlnGlnMetGlyThrGlnLysValIleAspArgLysValGlnLeuGlnLys 154
QY 479 GAGATGCTCTTGTACTTCACAAAGCGCTTAAGCCAGGCGGAGCCGAAATGGCTGCCAG 538
Db 155 GluAspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaLys 174
QY 539 CTGCATCCATGACAAAGCGAAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGAGA 598
Db 175 LeuHisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGlnGlnThrCyl 194
QY 599 ATCTCTCTAGTAATTTCTGCATGCTGATGACATGACAGAGAGGTGGTCCCTTGA 658
Db 195 IleSerHisSerGlnPheLeuHisArgGlnHisThrGlnGlnGlyGlnProLeuGly 214
QY 659 GGCTCTGAGATTTCTTAACCTTGCATCTTGGCTTGCATGAGCCAGCGCTGTA 718
Db 215 GlyProLeuLysPheSerAsnProCysIleLeuTrpLeuHisAlaAspGlnAlaValLeu 234
QY 719 GATGAGCGCTTGCGATTAAGAGCGTGCATGACATGCTTGCCTGGCGCTTGAGAGACTA 778
Db 235 AspGluArgLeuAspLysArgValAspAspMetLeuAlaIleGlyLeuGluGluLeu 254
QY 779 AGGATTTTCACAGACGCTTAATCAAGAAGATGTTTGGGAAATAGCCAGACTATCA 838
Db 255 ArgAspPheHisArgArgTyrAsnGlnLysAsnValSerGlnAsnSerGlnAspTyrGln 274
QY 839 CATGTAATCTCCAAATTCATTTGGCTTCAAGAAATTTTCAGAGTACTGATCACTAGAGGA 898
Db 275 HisGlyIlePheGlnSerIleGlyPheLysGluPheHisGluTyrLeuIleThrGlnGly 294
QY 899 AAATGCACACTGAGAGACTAGTAACACAGCTTCTTAAGAAGA 940
Db 295 LysCysThrLeuGluThrSerAsnGlnLeuLeuLysGly 308

```

RESULT 5

```

PCT-US01-01239-1394
; Sequence 1394, Application PC/TUS0101239
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT213PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01239
; PRIOR FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 2318
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1394
; LENGTH: 222
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01239-1394

Alignment Scores:
Pred. No.: 3,78e-112 Length: 222
Score: 123.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.61% Indels: 0
DB: 1 Gaps: 0

US-09-513-151-3 (1-2041) x PCT-US01-01239-1394 (1-222)

OY 2 TGCCATTAAGATGGCGTCCGCGGCGGCGACGACGAGTTCCTGTGGGCGAGTGGCTCAG 61
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
OY 62 GGCCTGCAACGACGACCTTACCTTGTAGTATCTCCGGGCGACGGGCAACGGCAATCC 121
Db 21 GlyLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40
OY 122 ACCGTGGCGTTGACGTAGCCACGCGCTCGCGGTAGATCGTACGCTGACTCATG 181
Db 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyIleValIleValSerAlaAspSerMet 60
OY 182 CAGTCTATGAGGCGCTAGCATCATCACCACCAAGTTTCTGCCCAAGAGCAGAGATC 241
Db 61 GlnValTyrGluGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80
OY 242 TGCCGCGACCATGATGATCCTTGTGATCCTCTTGACCAATTAACAGTGGTGAC 301
Db 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrThrValAsp 100
OY 302 TTCGAAATAGACCACTGCTCTGATTAAGATATATTGGCCGACGACAAATTCCTATT 361
Db 101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle 120
OY 362 GTTGCGGA 370
Db 121 ValValGly 123

RESULT 6
PCT-US01-01349-586
; Sequence 586, Application PC/TUS0101349
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01349
; CURRENT FILING DATE: 2001-01-14
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 586
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01349-586
```

```

Alignment Scores:
Pred. No.: 3,78e-112 Length: 222
Score: 123.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.61% Indels: 0
DB: 1 Gaps: 0

US-09-513-151-3 (1-2041) x PCT-US01-01349-586 (1-222)

OY 2 TGCCATTAAGATGGCGTCCGCGGCGGCGACGACGAGTTCCTGTGGGCGAGTGGCTCAG 61
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
OY 62 GGCCTGCAACGACGACCTTACCTTGTAGTATCTCCGGGCGACGGGCAACGGCAATCC 121
Db 21 GlyLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40
OY 122 ACCGTGGCGTTGACGTAGCCACGCGCTCGCGGTAGATCGTACGCTGACTCATG 181
Db 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyIleValIleValSerAlaAspSerMet 60
OY 182 CAGTCTATGAGGCGCTAGCATCATCACCACCAAGTTTCTGCCCAAGAGCAGAGATC 241
Db 61 GlnValTyrGluGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80
OY 242 TGCCGCGACCATGATGATCCTTGTGATCCTCTTGACCAATTAACAGTGGTGAC 301
Db 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrThrValAsp 100
OY 302 TTCGAAATAGACCACTGCTCTGATTAAGATATATTGGCCGACGACAAATTCCTATT 361
Db 101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle 120
OY 362 GTTGCGGA 370
Db 121 ValValGly 123

RESULT 7
PCT-US01-01349-784
; Sequence 784, Application PC/TUS0101349
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01349
; CURRENT FILING DATE: 2001-01-14
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 784
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01349-784

Alignment Scores:
Pred. No.: 3,78e-112 Length: 222
Score: 123.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.61% Indels: 0
DB: 1 Gaps: 0

US-09-513-151-3 (1-2041) x PCT-US01-01349-784 (1-222)
```

Oy	4	TGCATTAAGATGGGCTCCGTGGCGGCTGCACGACAGACTCTGTGGAGCAATGGGCTGACG	61
Db	1	CysHsIysMetAlaSerValAlAlaAlaIarGhIaValaProValGlySerGlyIleuauir	20
Oy	62	GGCGTGCACAGGACCCTAACCTCTGTAGTATTTCTCGGGGCGACGGGACCGGCAAAATCC	121
Db	21	GlyIeuclnIarGhrIleuProIeuValIaValIleuclIyalarIrrGlyThrGlyLysSer	40
Oy	122	ACGCTGCGCTTGACAGCTAGGCGACAGCGCTGGCGGTGAGATCCTCACCAGCTGACTCCATG	18

QY	62	GCCTCGCAACGAGACCCCTACGCTCTTGTACTGATTCCTGGGGCCACGGGAGACCGGCCAATCC	12
Db	21	GlyLeuGlnIaArgThrLeuProLeuValValIleLeuGlnGlyAlaThrInGlyThGlyLysSer	40
QY	122	ACGCTGGCGCTTGCACACTAGGCGACGCGCTCGCGGTGAGATCGTACAGCGTGCATCCATG	181
Db	41	ThrLeuAlaLeuGlnIleuGlnGlyIaArgLeuGlnGlyGluIleValSerAlaIaAspSerMet	60
QY	182	CAGGTCTATGAAGCGCTAGACATCATATACCAACAAGGTTCTTGGCCCAAGACAGACAAATC	241
Db	61	GlnValTyrGlnGlyLeuAspPheIleIleThrAsnLysValSerAlaGlnGlnGlnIaArgIle	80
QY	242	TGCCGGACCAACATGATCAGCTTTGTGGATCTCTTGTGACAAATTATCACAGAGGTGGAC	301
Db	81	CysArgIshHisMetIleSerPheValIaAspProLeuValIleIhrAsnTyrThrValIaValAsp	100


```

OY      302 TTGGAANTAGCACCTCCTGCTGATTTGAAGATATATTTGCCCGGACAAATTCGATT 361
        |||||||
Db      101 PheArgSerArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle 120
OY      362 GTTGTGGCA 370
        |||||||
Db      121 ValValCys 123

RESULT 10
US-09-764-902-1394
: Sequence 1394, Application US/09764902
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT213
: CURRENT APPLICATION NUMBER: US/09/764,902
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - refer to PALM or file wrapper
: NUMBER OF SEQ ID NOS: 2318
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1394
: LENGTH: 222
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (124)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (145)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-902-1394

```

Alignment Scores:			
Pred. No.:	3 78e-112	Length:	222
Score:	123.00	Matches:	123
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	18.61%	Indels:	0
DB:	21	Gaps:	0
US-09-513-151-3 (1-2041) x US-09-764-902-1394 (1-222)			
QY	2 TGGCATTAGATGGCGGCTCCGTCGGGAGCTGCACGACAGTTCTCTGTGGCGAGTGGCCTCAGG	61	
DB	1 CysHSLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg	20	
QY	62 GGCCCTCAACGAGACCCCTACTCTTGTAAGTATTCCTCGGGCCACGGGACCCGCGCAATCC	121	
DB	21 GlyLeuGlnArgThrLeuProLeuValAlaIleLeuGlyAlaThrGlyThrGlyLysSer	40	
QY	122 ACGCTGGGCTTGCACCTAGGCGACGCGCTCGGCGGTGAGATGTCAGCGCTGACTCATG	181	
DB	41 ThrLeuAlaLeuGlnLeuGlyGlnArgGlyGlyGluIleValSerAlaAspSerMet	60	
QY	182 CAGGTCTTGAAGGCTTGACATCATCCACCAAGGTTTGTGCCCAAGACAAAGATC	24	
DB	61 GlnValTYrGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle	80	
QY	242 TGCCGGCACCATGATAGCTTTGTGTGATCTCTTGACCAATTACACAGTGTGGAC	301	
DB	81 CysArgHisHisMetIleSerPheValAspProLeuValIThrAsnTYrThrValValAsp	100	
QY	302 TTCACAATAGAGCACTGCTCTGATTCAAGATATATTGGCCAGACAAATTCATTT	361	
DB	101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle	120	
QY	362 GTTGTGGGA	370	
DB	121 ValValGly	123	
RESULT 11			
US-10-072-326-586			

```

Sequence 586, Application US/10072326
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PJ206C1
CURRENT APPLICATION NUMBER: US/10/072,326
CURRENT FILING DATE: 2002-02-11.
Prior Application removed - See File wrapper or Palm
NUMBER OF SEQ ID NOS: 939
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 586
LENGTH: 222
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (114)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: misc_feature
LOCATION: (145)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-072-326-586

```

```

Alignment Scores:
Pred. No.: 3,78e-112 length: 222
Score: 123.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.61% Indels: 0
DB: 24 Caps: 0

US-09-513-151-3 (1-2041) x US-10-072-326-586 (1-222)

QY 2 TGCCATAAGATGGCGTCGCGGCGGCTGCACAGACAGTTCCTGTGGGCGATGGGCTCAGG 61
      |||||
Db 1 CysHs1LysMetAlaSerValAlAlAlaIaIaArgAlaValAlProValGlySerGlyLeuArg 20
QY 62 GGCCTGCACAGCAGCCCTACCTCTGTGTACTGATTCCTCGGGGCGACGGGACCCGGCAAAATCC 122
      |||||
Db 21 GlyLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaIaIaTrgIlyThrGlyLysSer 40
QY 122 ACGCTGCGGCTTGCAGCTAGCGGCGAGCGGCTCGCGGCTGAGATCGTCAGCGCTGACTCCATG 181
      |||||
Db 41 ThrLeuAlaLeuGlnIleuGlyGlnArgIleuGlyGlyIleValIleSerAlaIaIaSerMet 60
QY 182 CAGGCTATGAAGGCTTGCAGATCATCATCAACAAGGTTTGTGCCCAAGACAGAGATC 244
      |||||
Db 61 GlnValTYrGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnGlnArgIle 80
QY 242 TGCCGGCACCCACATGATGACGTTGTGGATGCTCTTGAGCAGCAATTTACAGTGGTGGAC 303
      |||||
Db 81 CysAlaGlnHsIaIaSerPheValAspProLeuValIaIaAsnTYrThrValValAsp 100
QY 302 TTCAGAAATAGACAACTGCTCTGATTGAATATATTGTGCCGACAGCAAAATTCCTGATT 363
      |||||
Db 101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120
QY 362 GTTGCGGA 370
      |||||
Db 121 ValValGly 123

RESULT 12
US-10-072-326-784
; Sequence 784, Application US/10072326
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1206C1
; CURRENT APPLICATION NUMBER: US/10/072,326
; CURRENT FILING DATE: 2002-02-11.
; Prior Application removed - See File wrapper or Palm
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 3.1

```



```

OY 1061 ACAGCCACTCCCAATAAAGTGCATTCAGTGAAGCGAGACAGCAAGAGAACTTTCACCTG 1120
      |||||||
Db 53 ThrIatIhPrIolIeYsIePrIotYraSnGluAlaGluIaSnIySaYsSerTYrHISLeu 72
      |||||||
OY 1121 TGTGACCTCTGTGATTCGAATTCATTCATTTGGGATCCGGAATGGCAGCGCACATPAAATCC 1180
      |||||||
Db 73 CysaSpLeuCySaSpaRgIleIleIleIleGlyAsPaRgIuTrpAlaIaHISIleYsSer 92
      |||||||
OY 1181 AAATCCACTTGAACCAACTGAGAAAGAAAGAGAAGATTGGACTTCAGTGTCTGCACACCC 1240
      |||||||
Db 93 LysSerHISLeuASnGlnIleuIySLySaRgARgARGLeuAsPeRsaPaIaValaISnThr 112
      |||||||
Db 113 IlegIuSerGlnSerValSeIProasp 121
      |||||||
RESULT 15
PCT-US01-00663-27779
: Sequence 27779, Application PC/TUS0100663
: GENERAL INFORMATION:
: APPLICANT: Molecular Dynamics, Inc.
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: PB 0004 WO 7
: CURRENT APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 04 February 2000 (04.02.00)
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 26 May 2000 (26.05.00)
: PRIOR APPLICATION NUMBER: US 09/632,365
: PRIOR FILING DATE: 03 August 2000 (03.08.00)
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 03 October 2000 (03.10.00)
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 27 September 2000 (27.09.00)
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 21 September 2000 (21.09.00)
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 30 June 2000 (30.06.00)
: NUMBER OF SEQ ID NOS: 38837
: SOFTWARE: Molecular Dynamics Sequence Listing Engine
: SEQ ID NO 27779
: LENGTH: 57
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AL033527.25
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
: OTHER INFORMATION: EST_HUMAN HIT: BE242161.1, EVALUE 3.00e-21
: OTHER INFORMATION: SWISSPROT HIT: O9KAC3, EVALUE 2.00e-04
PCT-US01-00663-27779

Alignment Scores:
Pred. No.: 1,29e-37 Length: 57
Score: 48.00 Matches: 48
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7,26% Indels: 0
DB: 1 Gaps: 0

US-09-513-151-3 (1-2041) x PCT-US01-00663-27779 (1-57)
OY 185 GTCTATGAGAGCCTAGACATCATCACCACAGAGTTTCTGCCAAGACAGACAGATCTGC 244
      |||||||
Db 1 ValTYcIuGcIyLeuASnSpIleIleThraSnIySaIeRaIaGlnGluIaRgIleGys 20
      |||||||
OY 245 CGGCAACCATGATGATACGCTTGTGATCCTCTGTGTGACCAATTACACAGTGTGCATTC 304
      |||||||

```

Db 21 ArgHisHisMetLeerPheValAspProLeuValThrIasnGlyThrValValAspPhe 40

Oy 305 AGAATAGAGCAACTGCTTCATT 328

Db 41 ArgAsnArgAlaThrAlaLeuLe 48

Search completed: April 21, 2003, 19:32:51
Job time : 280 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 18:59:06 ; Search time 86.5 seconds

(without alignments)
9723.518 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 661

Sequence: 1 CTGCCATAGATGCGTCGCG.....TTTACAGAGAAAAA 2041

Scoring table:

OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 671580 seqs, 206047115 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1343044

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame_plus_n2p.model -DEV=xlh
-Q/cgn2.1/USPTO-spool/US09513151/runat_15042003_141246.27221/app.query.fasta_1.2183
-DB=SPTRMBL_21 -QFMT=fasta -SUFFIX=olin2p.rspc -MIMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -STAR=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MIMLEN=0 -MAXLEN=200000000
-USER=US09513151.ecgn.1.1.112 @runat_15042003_141246.27221 -NCPU=3
-NO_XLPHY -NO_MAP -LARGESOURCE -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6 -Fgapext=7
-Ygapop=60 -Ygapext=60 -Delop=6 -Delext=7

Database :

1: SPTRMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_rvirus:*
17: sp_bacteriap:*
17: sp_archaeap:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	419	63.4	435	4	Q96L45
					Q96L45 homo sapien

ALIGNMENTS

RESULT 1	ID	PRELIMINARY	PRT:	435 AA.	
Q96L45	Q96L45	01-DEC-2001 (TREMBLrel. 19, Created)			Q9h3h1 homo sapien
AC	Q96L45	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			Q9nxt7 homo sapien
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				Q96fj3 homo sapien
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				Q9d1h5 mus musculus
DE	tRNA isopentenyl transferase (Fragment).				Q9r1d5 pseudomonas
OS	Homo sapiens (Human).				Q51532 pseudomonas
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				Q9ut75 schizosach
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				Q9qy66 oryza sativ
OX	NCBI_TaxId=9606;				Q9f525 mus musculus
RN	(1)				Q9f525 escherichia
RP	SEQUENCE FROM N.A.				Q9p1x9 campylobact
RX	MEDLINE=21444833; PubMed=11560893;				Q58821 pyrococcus
RA	Lemieux J., Lakowski B., Webb A., Meng Y., Ubach A., Bussiere F.,				Q13503 pleurotus o
RT	Barnes T., Hekimi S.,				Q9twq7 deinoococcus
RT	"Regulation of Physiological Rates in Caenorhabditis elegans by a				Q9kyh0 streptomyce
RT	tRNA-Modifying Enzyme in the Mitochondria.";				Q9f532 escherichia
RL	Genetics 159:147-157(2001).				Q52835 rhizobium 1
DR	EMBL: AY052768; AAL14107.1;				Q46323 graciliaria
DR	InterPro: IPR002627; IPT.				Q8wak7 neozelebori
DR	InterPro: IPR000822; Znf_C2H2.				Q9y9c6 aeropyrum p
DR	Pfam: PF01715; IPT; 1.				Q924b5 rattus norv
DR	ProDom: PD004674; IPT; 1.				Q9k798 bacillus ha
					Q9cpq3 mus musculu
					Q8velf mus musculu
					Q9cx70 mus musculu
					Q8x561 raltostonia s
					Q8u482 pyrococcus
					Q9svf5 arabidopsis
					Q9bx01 homo sapien
					Q33762 allomyces m
					Q8vf5 mus musculu
					Q9caa2 arabidopsis
					Q94id2 arabidopsis
					Q94ic9 arabidopsis
					Q94yb1 arabidopsis
					Q8velf mus musculu
					Q9wcd9 arabidopsis
					Q9vrg7 drosophila
					Q9vcs3 arabidopsis
					Q94id3 arabidopsis
					Q05699 nicotiana t
					Q9pbd0 xyella las
					Q8wbm0 matres zibe

DR SMART: SM00355; ZNF_C2H2.1;
DR TIGR00174; miaa:1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Transferase.
FT NON_TER 1
SO SEQUENCE 435 AA; 48948 MW; 2279AE7C2D999FE1 CRC64;

Alignment Scores:

Pred. No.:	0	Length:	435
Score:	419.00	Matches:	419
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	63.39%	Indels:	0
DB:	4	Gaps:	0

US-09-513-151-3 (1-2041) x Q96L45 (1-435)

```

QY 29 GCACGACGATTCCTGTGGCGAGTGGCGCTCAGAGGCGCTGCAACGACGACCTTCTGTGA 88
Db 1 AlaargAlaValProvalGlySerGlyLeuArgGlyLeuGlnArgThrLeuProLeuVal 20
QY 89 GTGATTCCTGGGGCCCGCCAGCCGCAATTCACGCTGGCGCTTGACAGCCAGCGG 148
Db 21 ValIleLeuGlyAlaThrGlyThrGlyLysSerThrLeuAlaLeuGlnLeuGlyGlnArg 40
QY 149 CTGGGGCGTGAGATCGTCGACGCTGACATGCAAGCTGTATGGAAGCGCTAGACATCATC 208
Db 41 LeuGlyGlyGlnIleValSerAlaAspSerMetGlnValTyrGlyGlyLeuAspIleLe 60
QY 209 ACCAACAAGCTTTCTGCCCAAGACAGCAGAGATCTGCCGCGACCATGATCAGCTTTGTG 268
Db 61 ThrAsnLysValSerAlaGlnGlnGlnArgIleCysArgGlnHisMetIleSerPheVal 80
QY 269 GATCCTCTGTGGACCAATTCACAGTGGTGGACTTCAGAAATAGACAACTGCTGTGATT 328
Db 81 AspProLeuValThrAspTyrThrValAlaSprPheArgAsnArgAlaThrAlaLeuIle 100
QY 329 GAAGTATATTGTGCCCGGACAAATTCCTATTGTGTGGGAGGAGCAACCAATTAATTAATT 388
Db 101 GlnAspIlePheAlaArgAspLysIleProIleValValGlyGlyThrAsnTyrTyrIle 120
QY 389 GAATCTCTGCTGTGGAAAGTCTTGTCAATACCAAGCCCGACAGATGGGACATGAGAA 448
Db 121 GluSerLeuLeuThrPlysValLeuValAsnThrLysProGlnGlnMetGlyThrGlnLys 140
QY 449 GTGATTCACCGAAAGTGGACCTTGAAGAAGAGATGCTGTGTAATTCACAAACGCTTA 508
Db 141 ValIleAspArgLysValGlyIleLeuGlnLysGlnAspGlyLeuValIleGlnHisLysArgLeu 160
QY 509 AGCCAGGCGGACCCAGAAATGGCTGCCAAGCTGCATCCACATGCAACGCAAAAGTGGCC 568
Db 161 SerGlnValAspProGlnMetAlaAlaLysLeuHisProHisAspLysAlaGlyValAla 180
QY 569 AGGAGCTTGCAGATTTTGAAGAAACAGAAATCTCTCATAGTGAATTTCTCCATGCTCAA 628
Db 181 ArgSerLeuGlnValPheGlnGlnThrGlyLysSerHisSerGlnPheLeuHisArgGln 200
QY 629 CATACGCAAGAGAGTGGTGGCTCCCTTGGAGGCTCTGGAAGTTCTCTAACCCCTGCATC 688
Db 201 HisThrGlnGlnGlyGlyGlyProLeuGlnGlyProLeuLysPheSerAsnProCysIle 220
QY 689 CTTGGGCTTCATGCTGACGACGAGTCTATGATGACGCGCTGGATAGAGAGGTGGATAC 748
Db 221 LeuThrPheHisAlaAspGlnAlaValLeuAspGlnAlaGlyLeuAspLysAlaGlyValAsp 240
QY 749 ATGCTTGCTGCTGGGCTCTTGGAGAGAACTAAGATTTTCCACAGCGCTAATACAGAG 808
Db 241 MetLeuAlaAlaGlyLeuLeuGlnGlnGlnLeuArgAspPheHisArgGlyTyrAsnGlnLys 260
QY 809 AATGTTTCGGAATACCGACGAGACTATACACATGCTATCTTCCATCAATATGGCTTCAAG 868
Db 261 AsnValSerGlnAsnSerGlnAspTyrGlnHisGlyIlePheGlnSerIleGlyPheLys 280

```

```

QY 869 GAATTCACGAGTACATGATCATGAGGAAATGCACACTGGAGACTAGTAACAGCTT 928
Db 281 GluPheHisGlyTyrLeuIleThrGlnGlyLysCysThrLeuGlnThrSerAsnGlnLeu 300
QY 929 CTAAGAAGACCTGTGCTCCATTTGCTCCCTGTGTATGCTTATAGAGATATCTGATGTC 988
Db 301 LeuLysLysGlyProGlyProIleValProProValTyrGlyLeuGlnValSerAspVal 320
QY 989 TCGAAGTGGAGAGACTGTGTTCTTGAACCTGCTCTTGAATCGTGCAAAGTTTATCCAG 1048
Db 321 SerLysTrpGlnGlnSerValLeuGlnProAlaLeuGlnIleValGlnSerPheIleGln 340
QY 1049 GGCCACAAGCCCTACAGCCGACCTCCAAATAGAATGCCATCATGAAGCTGCAACAAAGA 1108
Db 341 GlnHisLysProThrAlaThrProIleLysMetProTyrAsnGlnAlaGlnLysArg 360
QY 1109 AGTTATCACCTGTGTGACCTGTGTATGATGCAATTCATATGCGGATGCCGAATGGCAGCC 1168
Db 361 SerTyrHisLeuCysAspLeuCysAspArgIleIleIleGlyAspArgGlnTrrAlaAla 380
QY 1169 CACATTAATCCAAATCCCACTTGACCACTGAGAAAGAAAGAGATTGGACTCAGAT 1228
Db 381 HisIleLysSerLysSerHisLeuAsnGlnLeuLysLysArgArgLeuAspSerAsp 400
QY 1229 GCTGTCAACACCATAGAAAGTCAGAGTGTCCCGACGCTATTACAAAGACCTAAA 1285
Db 401 AlaValAsnThrIleGluSerGlnSerValSerProAspTyrAsnLysGlnProLys 419

```

RESULT 2

Q9H3H1 PRELIMINARY; PRT; 467 AA.

```

AC Q9H3H1:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE tRNA isopentenyl pyrophosphate transferase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20564178; PubMed=11111046;
RA Golovko A., Hjaln G., Sibson F., Nicander B.;
RT "Cloning of a human tRNA isopentenyl transferase.";
FL Gene 258:85-93(2000).
DR EMBL: AF074918; AAC31324.1;
DR InterPro: IPR002627; IPT.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF01715; IPT.1.
DR ProDom: PD004674; IPT.1.
DR TIGR00174; miaa:1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 467 AA; 52725 MW; 634469919D7F56A5 CRC64;

```

Alignment Scores:

Pred. No.:	0	Length:	467
Score:	310.00	Matches:	310
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	46.90%	Indels:	0
DB:	4	Gaps:	0

US-09-513-151-3 (1-2041) x Q9H3H1 (1-467)

```

QY 11 ATGCGTCGCGGGCGCTGCACGACGATTCCTGTGGCGAGTGGGCTCAGAGGCGCTGCA 70
Db 1 MetAlaSerValAlaAlaAlaAlaArgAlaValProValIleSerGlyLeuArgGlyLeuGln 20
QY 71 CGGACCTTACCTCTTGTAGTGATTTCTGGGGCCACGGGCAACCGGCAATTCACGCTGGCG 130
Db 21 ArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSerThrLeuAla 40

```

```

OY 131 TTGCAGCTAGGCGGCGGCGGTGAGTGCATGCGCTGATGCGAGCTAT 190
DB 41 LeuGlnLeuGlyGlnArgLeuGlyGlyGluIleValSerAlaAspSerMetGlnValTyr 60
OY 191 GAAGGCTGAGATCATCATCCAAAGAGCTTTCTGCCCAAGACAGAGAAATCTCCGCGAC 250
DB 61 GluGlyLeuAspIleIleThrAsnIlyValSerAlaGlnGlnGlnArgIleCysArgHis 80
OY 251 CACATGATCAGCTTTGTGATCCCTCTGTGACCAATATACACAGCTGGTGCATTCAGAAAT 310
DB 81 HisMetIleSerPheValAspProLeuValThrAsnTyrThrValAlaAspPheArgAsn 100
OY 311 AAGAGCACTGCTGATTAAGATATATTGCCCGAGACAAATCTCTTTGTGGGA 370
DB 101 ArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspIlySerProIleValGly 120
OY 371 GGAACCAATTTATACATTGAATCTCTCTGGAAGTTCTTGTCAATACAGACCCAC 430
DB 121 GlyThrAsnTyrTyrIleGluSerLeuLeuTyrPlyValLeuValAsnThrLysProGln 140
OY 431 GAGATGGGCACTGAGAAAGTATGACCCGAAAGTGCAGCTTGAAGAGAGATGCTT 490
DB 141 GluMetGlyThrGluLysValIleAspArgLysValGluGluGlyGluAspGlyLeu 160
OY 491 GTACTTCACAAAGCCCTAAGCCAGGTGGAGCCAGAAATGGCTGCAAGCTGCATCCCAT 550
DB 161 ValLeuHisLysArgLeuSerGlnValAspProGluMetAlaAlaLysIleHisProHis 180
OY 551 GACAAAGCGAAAGTGGCCAGAGCTTGCAAGTTTGAAGAAACAGAGAACTCTCATAGT 610
DB 181 AspLysArgLysValAlaArgSerLeuGlnValPheGluGluThrGlyIleSerHisSer 200
OY 611 GAATTTCCATGCTGATACATACAGGAAGAGGTGGTGGTCCCTTGGAGCTCTCTGAG 670
DB 201 GluPheLeuHisArgGlnHisThrGluGluGlyGlyGlyProLeuGlyGlyProLeuLys 220
OY 671 TTCTTAACCTTGCATCTTGGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 730
DB 221 PheSerAsnProCysIleLeuTyrPheLeuHisAlaAspIleAlaValIleAspGluArgLeu 240
OY 731 GATTAAGGCTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 790
DB 241 AspLysArgValAspAspMetLeuAlaAlaGlyLeuGluGluGluLeuArgAspPheHis 260
OY 791 AGAGCGTATATACAGAAATGTTTGGGAAATAGCCAGAGACTATCAACATGATCTTC 850
DB 261 ArgArgTyrAsnGlnLysAsnValSerGluAsnSerGlnAspTyrGlnHisGlyIlePhe 280
OY 851 CAATCATTTGGCTTCAAGGAATTTCAAGAGTACTGATCACTGAGGGAATTCACACTG 910
DB 281 GlnSerIleGlyPheLysGluPheHisGluTyrLeuIleThrGluGlyLysCysThrLeu 300
OY 911 GAGACTAGTAACAGCTTCTTAAGAAAGA 940
DB 301 GluThrSerAsnGlnLeuLeuLysGly 310

```

RESULT 3

```

O9NXT7 PRELIMINARY: PRT: 326 AA.
AC O9NXT7:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CDNA FLJ20061 fls, clone COL01383.
OS Homo sapiens (Human).
OC Eukaryota: Metazoa: Chordata: Craniala: Vertebrata: Euteleostomi:
OC Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Kawabata A., Hiki J. T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,

```

```

RA Okitani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isonai T., Sugano S.;
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK000068: BAA90923.1;
DR InterPro: IPR002627; IPRF.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF01715; IPRF. 1.
DR ProDom: PD004674; IPRF. 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2.1; UNKNOWN.1.
SQ SEQUENCE 326 AA, 37435 MW, EAA3F0F9664B7ACE CRC64:

```

Alignment Scores:

```

Pred. No.: 3,45e-171 Length: 326
Score: 169.00 Matches: 169
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25,578 Indels: 0
DB: Gaps: 4

```

US-09-513-151-3 (1-2041) x O9NXT7 (1-326)

```

OY 434 ATGGGCACTGAGAAAGTATGACCCGAAAGTGGAGCTTGAAGAGAGATGCTTTGTA 493
DB 1 MetGlyThrGluLysValIleAspArgLysValGluLeuGluLysGluAspGlyLeuVal 20
OY 494 CTTTCACAAAGCCCTAAGCCAGGTGGAGCCAGAAATGGCTGCAAGCTGCATCCCATGAC 553
DB 21 LeuHisLysArgLeuSerGlnValAspProGluMetAlaAlaLysIleHisProHisAsp 40
OY 554 AAAGCGAAGTGGCCAGAGCTTGCAAGTTTGAAGAAACAGAGAACTCTCATAGTGA 613
DB 41 LysArgLysValAlaArgSerLeuGlnValPheGluGluThrGlyIleSerHisSerGlu 60
OY 614 TTCTTCATGCTGATACATACAGGAAGAGGTGGTGGTCCCTTGGAGCTCTCTGAGTTC 673
DB 61 PheLeuHisArgGlnHisThrGluGluGlyGlyGlyProLeuGlyGlyProLeuLysPhe 80
OY 674 TCTTAACCTTGCATCTTGGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 733
DB 81 SerAsnProCysIleLeuTyrPheLeuHisAlaAspIleAlaValIleAspGluArgLeuAsp 100
OY 734 AAGAGGCTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
DB 101 LysArgValAspAspMetLeuAlaAlaGlyLeuLeuLeuGluLeuAlaArgAspPheHisArg 120
OY 794 CGCTATATACAGAAATGTTTGGGAAATAGCCAGAGACTATCAACATGATCTTCCAA 853
DB 121 ArgTyrAsnGlnLysAsnValSerGluAsnSerGlnAspTyrGlnHisGlyIlePheGln 140
OY 854 TCAATGCTTCAAGGAATTTCAAGAGTACTGATCACTGAGGGAATTCACACTGAG 913
DB 141 SerIleGlyPheLysGluPheHisGluTyrLeuIleThrGluGlyLysCysThrLeuGlu 160
OY 914 ACTAGTAACAGCTTCTTAAGAAAGA 940
DB 161 ThrSerAsnGlnLeuLeuLysGly 169

```

RESULT 4

```

O96FJ3 PRELIMINARY: PRT: 324 AA.
AC O96FJ3:
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Similar to tRNA isopentenylpyrophosphate transferase.
OS Homo sapiens (Human).
OC Eukaryota: Metazoa: Chordata: Craniala: Vertebrata: Euteleostomi:
OC Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;

```

RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC010741; AAH10741.1; -
 DR InterPro: IPR002627; IPRP.
 DR InterPro: IPR000822; Znf_C2H2.
 DR Pfam: PF01715; IPRP; 1.
 DR ProDom: PD004674; IPRP; 1.
 DR SMART: SM00355; Znf_C2H2; 1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 DR TRANSFERASE.
 KW SEQUENCE 324 AA; 37223 MW; 1E6835D7C09126A9 CRC64;

Alignment Scores:

Pred. No.:	1,13e-117	Length:	324
Score:	119.00	Matches:	167
Percent Similarity:	96.53%	Conservative:	0
Best Local Similarity:	96.53%	Mismatches:	0
Query Match:	18.00%	Indels:	6
DB:	4	Gaps:	0

US-09-513-151-3 (1-2041) x Q96FJ3 (1-324)

OY 434 ATGGCAGTGAAGTGAATGACCGAAGAGTGGAGCTTGAAGAAGAGAGATGCTTGTGA 493
 DB 1 MetGlyThrGluValIleAspArgLysValGluLeuGluLysGluAspGlyLeuVal 20
 OY 494 CTTCACAAAGCCCTAAGCCAGTGGACCCAGAAATGGCTGCCAGCTGCATGCATGAC 553
 DB 21 LeuHisLysArgLeuSerGlnValAspProGluMetAlaIleLysLeuHisProHisAsp 40
 OY 554 AAACCCAAAGTGGCCAGAGAGCTGCAAGCTTTTGAAGAAGCAAGATCTCATAGTGA 613
 DB 41 LysArgLysValAlaArgSerLeuGlnValPheGluGluThrGlyIleSerHisSerGlu 60
 OY 614 TTTCCTCCTGCAACATACGGAAGAGCTGGTCCCTTGGAGGCTCTGCAAGTTC 673
 DB 61 PheLeuHisArgGlnHisThrGluGluGlyGlyProLeuGlyGlyProLeuLysPhe 80
 OY 674 TCTAACCTTGATCCCTTGGCTTCATGCTGACAGAGATTTAGATGACGGCTTGAT 733
 DB 81 SerAsnProCysIleLeuThrPheHisAlaAspGlnAla-----AspGluArgLeuAsp 98
 OY 734 AAGAGCGGTGATGACATGCTGCTGCTGGCTGCTGGAAGCAAGATTTTACAGAGA 793
 DB 99 LysArgValAspAspMetLeuAlaIleGlyLeuLeuGluGluLeuAspAspPheHisArg 118
 OY 794 CGCTATATATCAGAGAAATGTTTGGAAATAGCCAGAGATTCACATGCTATCTTCAA 853
 DB 119 ArgTyrIasnGlnLysAsnValSerGluAsnSerGlnAspTyrGlnHisGlyIlePheGln 138
 OY 854 TCAATTGGCTTCAAGAAATTTACAGAGTACCTGATCAGGGAATGACAGCTGAG 913
 DB 139 SerIleGlyPheLysGluPheHisGlyTyrLeuIleThrGluGlyLysCysThrLeuGlu 158
 OY 914 ACTAGTACCAAGCTTTAAAGAAAGCA 940
 DB 159 ThrSerAsnGlnLeuLeuLysGly 167

RESULT 5
 Q9DIH5 PRELIMINARY: PRT: 326 AA.

AC Q9DIH5: 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE 2310075G14RLK protein (Riken cDNA 2310075G14 gene).
 GN 2310075G14RLK
 OS Mus musculus (Mouse).
 OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
 OC Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus:
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE-EMBRYO;
 RX MEDLINE-21085660; PubMed-11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Sakai K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
 RA Kuenl L.M., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., But C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzaroli J., Mombers P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K., F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN (2)
 RP SEQUENCE FROM N.A.

RC TISSUE-KIDNEY;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK003556; BAB2853.1; -
 DR EMBL: BC019812; AAH19812.1; -
 DR MGD: MGI:1914216; 2310075G14RLK.
 DR InterPro: IPR002627; IPRP.
 DR InterPro: IPR000822; Znf_C2H2.
 DR Pfam: PF01715; IPRP; 1.
 DR ProDom: PD004674; IPRP; 1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 DR SEQUENCE 326 AA; 37191 MW; 454367A8B70DD1F0 CRC64;

Alignment Scores:
 Pred. No.: 1.52e-39 Length: 326
 Score: 46.00 Matches: 46
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.96% Indels: 0
 DB: 11 Gaps: 0

US-09-513-151-3 (1-2041) x Q9DIH5 (1-326)

OY 494 CTTCACAAAGCCCTAAGCCAGTGGACCCAGAAATGCTGCCAAGCTGCATGCATGAC 553
 DB 21 LeuHisLysArgLeuSerGlnValAspProGluMetAlaIleLysLeuHisProHisAsp 40
 OY 554 AAACCCAAAGTGGCCAGAGAGCTGCAAGCTTTTGAAGAAGCAAGATCTCATAGTGA 613
 DB 41 LysArgLysValAlaArgSerLeuGlnValPheGluGluThrGlyIleSerHisSerGlu 60
 OY 614 TTTCCTCCTGCAACATACGGAAGAGCTGGTCCCTTGGAGGCTCTGCAAGTTC 673
 DB 61 PheLeuHisArgGlnHis 66

RESULT 6
 Q9RLD5 PRELIMINARY: PRT: 381 AA.

AC Q9RLD5: 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Putative secretion protein.
 GN P110.
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria: Proteobacteria: gamma subdivision: Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=316;
 RN (1)

RP SEQUENCE FROM N.A.
 RC STRAIN-JM300;
 RX MEDLINE-21359310; PubMed-11466271;
 RA Graupner S., Weger N., Sohn M., Wackernagel W.;
 RT "Requirement of bovel competence genes pilT and pilU of Pseudomonas
 stutzeri for natural transformation and suppression of pilT deficiency
 by a Hexahistidine tag on the type IV pilus protein PilA1.";
 RL J. Bacteriol. 183:4694-4701(2001).
 DR EMBL: AJ248385; CAB56296.1;
 DR InterPro: IPR001482; GSPIL_E.
 DR Pfam: PF00437; GSPIL_E.1.
 DR ProDom: PD000739; GSPIL_E.1.
 SQ SEQUENCE 381 AA; 42417 MW; C69B6E6BC24EA3DB CRC64;

Alignment Scores:
 Pred. No.: 5.94 Length: 381
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.36% Indels: 0
 DB: Gaps: 0

US-09-513-151-3 (1-2041) x Q9RLD5 (1-381)
 OY 98 GGGCCGACGGCAGCGCAATCCACG 124
 DB 130 GYAlaThrGlyThrGlyLysSerThr 138

RESULT 7
 O51532 PRELIMINARY; PRT; 382 AA.
 AC 051532;
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
 DE PilU (Twitching motility protein PilU).
 GN PilU OR PA0396.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE-91285432; PubMed-1676385;
 RA Whittechurch C.B., Hobbs M., Livingston S.P., Krishnapillai V.,
 RA Matlick J.S.;
 RT "Characterisation of a Pseudomonas aeruginosa twitching motility gene
 and evidence for a specialised protein export system widespread in
 eubacteria.";
 RL Gene 101:33-44(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RA Whittechurch C.B., Matlick J.S.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE-20437337; PubMed-10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yada Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig I.T.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL: L27667; AAA25965.1;
 DR EMBL: AE004477; AAG03785.1;
 DR InterPro: IPR001482; GSPIL_E.
 DR Pfam: PF00437; GSPIL_E.1.

DR ProDom: PD000739; GSPIL_E.1.
 KW Complete Proteome.
 SQ SEQUENCE 382 AA; 42532 MW; 02665D0FED35E17 CRC64;

Alignment Scores:
 Pred. No.: 5.94 Length: 382
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.36% Indels: 0
 DB: Gaps: 16

US-09-513-151-3 (1-2041) x Q51532 (1-382)
 OY 98 GGGCCGACGGCAGCGCAATCCACG 124
 DB 130 GYAlaThrGlyThrGlyLysSerThr 138

RESULT 8
 Q9UT75 PRELIMINARY; PRT; 434 AA.
 AC Q9UT75;
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE tRNA isopentenyltransferase.
 GN SPAC343.15.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OC NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972H;
 RA Murphy L., Harris D., Wood V., Rajandream M.A., Barrell B.G.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL109739; CAB52278.1;
 DR InterPro: IPR002627; IPPT.
 DR Pfam: PF01715; IPPT.1.
 DR ProDom: PD004674; IPPT.1.
 DR TrGFams: TRG00174; mlaa.1.
 DR tRNA transferase.
 KW tRNA transferase.
 SQ SEQUENCE 434 AA; 50121 MW; FCFD5101DF95F34D CRC64;

Alignment Scores:
 Pred. No.: 5.84 Length: 434
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.36% Indels: 0
 DB: Gaps: 3

US-09-513-151-3 (1-2041) x Q9UT75 (1-434)
 OY 350 AAAATTCATATTGTCGAGGACAC 376
 DB 93 LysIleProIleValIleGlyGlyThr 101

RESULT 9
 O8S926 PRELIMINARY; PRT; 439 AA.
 ID O8S926;
 AC O8S926;
 DT 01-JUN-2002 (TEMBLrel. 21, Created)
 DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE Similar to tRNA isopentenyltransferase.
 GN OJ1656.A11.14.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriarthoidae; Oryzaceae; Oryza.
 OC NCBI_TaxID=39947;

DR InterPro: IPR003509; UPF0102.
 DR Pfam: PF02021; UPF0102; 1.
 KW Hypothetical protein: Complete proteome.
 SQ SEQUENCE 112 AA; 13179 MW; 04EEAF317625A7DE CRC64;

Alignment Scores:

Pred. No.:	82.4	Length:	112
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	16	Gaps:	0

US-09-513-151-3 (1-2041) x O9P1X9 (1-112)

OY 573 GCCTGCAAGTTTTCAGAGAAACAG 596

Db 16 AlacystylsphenulstysGln 23

RESULT 13

058821 ID 058821 PRELIMINARY: PRT: 121 AA.

AC 058821: 01-AUG-1998 (TREMBlrel. 07, Created)

DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Hypothetical protein PH1094.

GN PH1094.

OS Pyrococcus horikoshii.

OC Archaea: Euryarchaeota: Thermococci: Thermococcales; Thermococcaceae;

OC Pyrococcus.

OX NCBI_Taxid=53953;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OT3;

RA MEDLINE=98344137; PubMed=9679194;

RA Kawarabayashi Y., Sawada M., Horikawa H., Hino Y.,

RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,

RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,

RA Funahashi T., Tanaka T., Kudon Y., Yamazaki J., Kushida N., Oguchi A.,

RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,

RA Masuchi Y., Shizuya H., Kikuchi H.;

RT "Complete sequence and gene organization of the genome of a hyper-

thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";

RL DNA Res. 5:55-76(1998)

DR EMBL: AP000004; BAA30193.1;

KW Hypothetical protein: Complete proteome.

SQ SEQUENCE 121 AA; 12859 MW; 1BE4A036401AE523 CRC64;

Alignment Scores:

Pred. No.:	81.5	Length:	121
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.22%	Indels:	0
DB:	17	Gaps:	0

US-09-513-151-3 (1-2041) x 058821 (1-121)

OY 1111 ACTGCTCTGTCTCAGCTTCATT 1088

Db 6 ThrserleuValleuSerPhele 13

RESULT 14

013503 ID 013503 PRELIMINARY: PRT: 131 AA.

AC 013503: 01-JAN-1998 (TREMBlrel. 05, Created)

DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)

DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE Hydrophobin precursor.

GN POH2.

OS Pleurotus ostreatus (oyster mushroom) (White-rot fungus).

OC Eukaryota: Fungi: Basidiomycota: Hymenomycetes: Homobasidiomycetes;
 OC Agaricales: Pleurotaceae: Pleurotus.

OX NCBI_Taxid=5322;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=FRUIT BODY;

RA Asgelsdottil S.A., de Vries O.M.H., Wessels J.G.H.;

RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=R7;

RA Asgelsdottil S.A., de Vries O.M.H., Wessels J.G.H.;

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: Y14657; CAA74987.1;

DR EMBL: AJ225061; CAA12392.1;

DR InterPro: IPR00338; Hydrophobin.

DR Pfam: PF01185; Hydrophobin; 1.

DR SMART: SM00075; HYDRO; 1.

DR PROSITE: PS00956; HYDROPHOBIN; 1.

KW Signal.

FT SIGNAL 1 23 POTENTIAL.

FT CHAIN 24 131 POTENTIAL.

SQ SEQUENCE 131 AA; 13114 MW; 739EA57E07FE42P1 CRC64;

Alignment Scores:

Pred. No.:	80.6	Length:	131
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.22%	Indels:	0
DB:	3	Gaps:	0

US-09-513-151-3 (1-2041) x 013503 (1-131)

OY 774 TCCTCCAGACCCAGCAGCAAC 751

Db 63 SerSerlySerProAlaAlaser 70

RESULT 15

09RM07 ID 09RM07 PRELIMINARY: PRT: 162 AA.

AC 09RM07: 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Hypothetical protein DR0609.

GN DR0609.

OS Deinococcus radiodurans.

OC Bacteria: Thermus/Deinococcus group: Deinococci: Deinococcaceae;

OC Deinococcaceae: Deinococcus.

OX NCBI_Taxid=1299;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=R1;

RA MEDLINE=20036896; PubMed=10567266;

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,

RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,

RA Vamathevan J.S., Lam P., McDonald L., Ufferback T., Zaleski C.,

RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,

RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,

RT Fraser C.M.;

RT "Genome sequence of the radioresistant bacterium Deinococcus

radiodurans R1.";

RL Science 286:1571-1577(1999).

DR EMBL: AE001919; AAF10196.1;

DR TIGR: DR0609;

KW Hypothetical protein: Complete proteome.

SQ SEQUENCE 162 AA; 16942 MW; 9EE78E844F112C12 CRC64;

Alignment Scores:

Pred. No.:	78.4	Length:	162
Score:	8.00	Matches:	8

Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.21%
 DB: 16
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-09-513-151-3 (1-2041) x Q9RM07 (1-162)

QY 107 GGCACCGGCAATCCACGCTGCGC 130
 |||||
 Db 84 GlyThrGlySerThrLeuAla 91

Search completed: April 21, 2003, 19:18:35
 Job time : 94.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 18:55:45 ; Search time 23.5 Seconds

(without alignments)
7204.526 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 661

Sequence: 1 CTGCATAGATGGCGTCG.....TTTACAGAAAAA 2041

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 segs, 41476328 residues

Word size: 1

Total number of hits satisfying chosen parameters: 225614

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlh
-O=/c9n2.1/USPTO-spool/US09513151/runatc.15042003.141246.27211/app.query.fasta.1.2183
-DB=SwissProt_40 -QPM=fastan -SUFFIX=olind2p.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Dits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=oto -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09513151.ecgn.1.1.22.6runatc.15042003.141246.27211 -NCPU=6 -TCU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-NARN_TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	1.4	196	1 PSAD_CHLRE	039615 Chlamydomon
2	9	1.4	428	1 MODS_YEAST	P07884 saccharomyc
3	9	1.4	698	1 UVRB_MYCLE	P57891 mycobacteri
4	9	1.4	698	1 UVRB_MYCTU	006150 mycobacteri
5	8	1.2	92	1 RL37_EMENT	09c01 emericella
6	8	1.2	125	1 DIVC_BACSU	P37471 bacillus su
7	8	1.2	213	1 VARD_SULSO	09u49 sulfolobus
8	8	1.2	216	1 KCY_CHLMU	09p10 chlamydia m
9	8	1.2	216	1 KCY_CHLTR	08458 chlamydia t
10	8	1.2	216	1 KIRY_STRCO	050519 streptomyc
11	8	1.2	228	1 YTOB_ERWHE	047826 erwina her
12	8	1.2	240	1 MTRA_METMA	059640 methanosarc
13	8	1.2	264	1 NUDC_HAEIN	P44710 haemophilus
14	8	1.2	378	1 PLCD_HUMAN	09n725 homo sapien
15	8	1.2	417	1 YP28_CAEEL	P98061 caenorhabd
16	8	1.2	498	1 YK04_MYCTU	010852 mycobacteri
17	8	1.2	580	1 NPLA_YEAST	P33755 saccharomyc
18	8	1.2	656	1 UVRB_MYCGE	P47319 mycoplasma

19	8	1.2	657	1 UVRB_BACCA	P56981 bacillus ca
20	8	1.2	657	1 UVRB_MYCPN	P75558 mycoplasma
21	8	1.2	658	1 UVRB_LISIN	092844 listeria in
22	8	1.2	658	1 UVRB_LISMO	08445 listeria mo
23	8	1.2	660	1 UVRB_BACHD	09k69 bacillus ha
24	8	1.2	661	1 UVRB_BACSU	P37954 bacillus su
25	8	1.2	661	1 UVRB_STRAAM	099v17 staphylococ
26	8	1.2	663	1 UVRB_STRPY	099v17 staphylococ
27	8	1.2	665	1 UVRB_ANASP	08vxs9 anabaena sp
28	8	1.2	665	1 UVRB_THETH	056243 themus the
29	8	1.2	666	1 UVRB_UNEPA	09p724 ureaplasma
30	8	1.2	669	1 UVRB_SYNY3	055170 synechocyst
31	8	1.2	709	1 UVRB_MYCTU	P10125 micrococci
32	8	1.2	954	1 M3KA_HUMAN	002779 homo sapien
33	8	1.2	1093	1 YKDS_CAEEL	003563 caenorhabd
34	8	1.2	1304	1 S3B1_HUMAN	075533 homo sapien
35	8	1.2	1304	1 S3B1_MOUSE	099n69 mus musculi
36	8	1.2	1307	1 S3B1_XENLA	057683 xenopus lae
37	8	1.2	1557	1 DVAL1_DICVI	024702 dictyocaulu
38	7	1.1	38	1 EST5_DROMO	P10095 drosophila
39	7	1.1	62	1 SECE_SULSO	P58191 sulfolobus
40	7	1.1	98	1 ACYO_CHICK	P07032 gallus gall
41	7	1.1	98	1 YD22_MYCTU	010635 mycobacteri
42	7	1.1	106	1 Y270_SYNY3	P73889 synechocyst
43	7	1.1	108	1 R33A_SCHPO	09usx4 schizosacch
44	7	1.1	108	1 R33B_SCHPO	09usg6 schizosacch
45	7	1.1	108	1 Y253_THEMA	09w95 thermotoga

ALIGNMENTS

RESULT 1
PSAD_CHLRE
ID PSAD_CHLRE STANDARD: PRT: 196 AA.
AC 039615;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem I reaction center subunit II, chloroplast precursor
DE (Photosystem I 20 kDa subunit) (Pst-D).
GN PSAD.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RX MEDLINE=95288386; PubMed=7770547;
RA Farrah J.A., Frank G., Zuber H., Rochaix J.-D.;
RT "Cloning and sequencing of a cDNA clone encoding the photosystem I
RT psad subunit from Chlamydomonas reinhardtii.";
RL Plant Physiol. 107:1485-1486(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CC-406;
RX MEDLINE=97033539; PubMed=8879236;
RA Hahn D., Bemmoun P., Kueck U.;
RT "Altered expression of nuclear genes encoding chloroplast
RT polypeptides in non-photosynthetic mutants of Chlamydomonas
RT reinhardtii: evidence for post-transcriptional regulation.";
Mol. Gen. Genet. 252:362-370(1996).
CC - FUNCTION: PSAD CAN FORM COMPLEXES WITH FERREDOXIN AND FERREDOXIN-
CC OXIDOREDUCTASE IN PHOTOSYSTEM I (PS I) REACTION CENTRE. PSAD MAY
CC ENCODE THE FERREDOXIN-DOCKING PROTEIN. PSAD MAY
CC - SUBCELLULAR LOCATION: ASSOCIATED WITH STROMAL SIDE OF THE
CC - SIMILARITY: BELONGS TO THE PSAD FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: X79674; CA561422.1; -

DR EMBL: X74419; CA52440.1; -

DR InterPro: IPR003685; Psad.

DR Pfam: PF02531; Psad; 1.

KW Photosynthesis; Photosystem I; Chloroplast; Transit peptide;

KW Thylakoid; Membrane.

FT TRANSIT 1 35 CHLOROPLAST (POTENTIAL).

FT CHAIN 36 196 PHOTOSYSTEM I REACTION CENTER SUBUNIT II.

SO SEQUENCE 196 AA; 21342 MW; 3A30F8527A896FC6 CRC64;

Alignment Scores:

Pred. No.:	1.71	Length:	196
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.36%	Indels:	0
DB:	1	Gaps:	0

US-09-513-151-3 (1-2041) x PSAD-CHLRE (1-196)

OY 130 GTTCAGCTAGCCAGCGCTCGCGG 156

Db 21 VALAIAIAIARGPROAIAIAIARGARG 29

RESULT 2

MOD5_YEAST

ID MOD5_YEAST STANDARD: PRT: 428 AA.

AC P07884; Q12203;

DT 01-AUG-1988 (Rel. 08, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE tRNA isopentenyltransferase (EC 2.5.1.8) (isopentenyl-diphosphate: tRNA isopentenyltransferase) (IPTase) (IPTase) (IPTase).

GN MOD5 OR YOR274W.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=87172703; PubMed=3031457;

RA Natarian D., Dihanich M.E., Natarian N.C., Hopper A.K.;

RT "DNA sequence and transcript mapping of MOD5: features of the 5' region which suggest two translational starts.";

RL Mol. Cell. Biol. 7:185-191(1987).

RN (2)

RP SEQUENCE FROM N.A.

RC STRAIN=S288C;

RX MEDLINE=97051594; PubMed=8896271;

RA Cheret G., Bernardi A., Sor F.J.;

RT "DNA sequence analysis of the VP1-SNF2 region on chromosome XV of Saccharomyces cerevisiae.";

RL Yeast 12:1059-1064(1996).

RN (3)

RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.

RX MEDLINE=92052176; PubMed=1946403;

RA Slusher L.B., Gillman E.C., Martin N.C., Hopper A.K.;

RT "mRNA leader length and initiation codon context determine alternative AUG selection for the yeast gene MOD5.";

RL Proc. Natl. Acad. Sci. U.S.A. 88:9789-9793(1991).

RN (4)

RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.

RX MEDLINE=91203856; PubMed=1850093;

RA Gillman E.C., Slusher L.B., Martin N.C., Hopper A.K.;

RT "MOD5 translation initiation sites determine N6-isopentenyladenosine modification of mitochondrial and cytoplasmic tRNA.";

RL Mol. Cell. Biol. 11:2382-2390(1991).

RN (5)

RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.

RX MEDLINE=94187700; PubMed=8139535;

RA Boguta M., Hunter L.A., Shen W.C., Gillman E.C., Martin N.C., Hopper A.K.;

RT "Subcellular locations of MOD5 proteins: mapping of sequences sufficient for targeting to mitochondria and demonstration that RT mitochondrial and nuclear isoforms commingle in the cytosol.";

RL Mol. Cell. Biol. 14:2298-2306(1994).

CC - FUNCTION: RESPONSIBLE FOR THE MODIFICATION OF A37 TO ISOPENTENYL A37 OF BOTH CYTOSOLIC AND MITOCHONDRIAL TRNAS.

CC - CATALYTIC ACTIVITY: isopentenyl diphosphate + tRNA = diphosphate + tRNA containing 6-isopentenyladenosine.

CC - PATHWAY: BIOSYNTHESIS OF THE MODIFIED BASE ISOPENTENYLADENOSINE IN TRNAS.

CC - SUBCELLULAR LOCATION: MITOCHONDRIAL, CYTOPLASMIC AND NUCLEAR.

CC - ALTERNATIVE PRODUCTS: TWO ISOFORMS MAY BE PRODUCED FROM THE USE OF ALTERNATIVE INITIATION CODONS IN THE SAME READING FRAME.

CC - SIMILARITY: BELONGS TO THE IPT TRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M15991; AAA34785.1; -

DR EMBL: X89633; CA61780.1; -

DR EMBL: Z75182; CA99499.1; -

DR PIR: A26717; A26717.

DR SCD: S0005800; MOD5.

DR InterPro: IPR002627; IPTP.

DR Pfam: PF01715; IPTP; 1.

DR ProDom: PD004674; IPTP; 1.

DR TrnFRNAs: TrnR00174; mlaa; 1.

KW Transferase; tRNA processing; ATP-binding; Alternative initiation; Mitochondrion; Nuclear protein.

FT FT 1 428

FT CHAIN 12 428

FT INIT_MET 12 12

FT NP_BIND 21 28

FT DOMAIN 210 232

FT CONFLICT 313 313 MISSING (IN REF. 1).

FT CONFLICT 375 375 C->R (IN REF. 1).

SO SEQUENCE 428 AA; 50236 MW; A956B17ABC05161F CRC64;

Alignment Scores:

Pred. No.:	1.54	Length:	428
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.36%	Indels:	0
DB:	1	Gaps:	0

US-09-513-151-3 (1-2041) x MOD5_YEAST (1-428)

OY 350 AAATTCCTATGTGTGGAGGAACC 376

Db 104 LysleProllevalvalGlyGlyThr 112

RESULT 3

UVRB_MYCLE

ID UVRB_MYCLE STANDARD: PRT: 698 AA.

AC P57991;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Excinuclease ABC subunit B.

GN UVRB OR ML1387.

OS Mycobacterium leprae.

OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;

```

OC Actinomycetales; Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Elgmeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RA "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRB STIMULATES
CC THE APASE ACTIVITY OF UVRB IN THE PRESENCE OF UV-IRRADIATED
CC DOUBLE-STRANDED DNA. IT ALSO ENHANCES THE ABILITY OF UVRB TO BIND
CC TO UV-IRRADIATED DUPLEX DNA (BY SIMILARITY).
CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVRB, UVRB AND UVRB.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE UVRB FAMILY.
CC -1- SIMILARITY: CONTAINS 1 UVR DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL583921; CAC31768.1;
DR HSSP: P56981; 1D9X.
DR Leproma; ML1387;
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR004807; UVRB.
DR InterPro: IPR001943; UVRB/C.
DR Pfam: PF00270; DEAD; 1.
DR Pfam: PF00271; helicase_C; 1.
DR Pfam: PF02151; UVR; 1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELIC; 1.
DR TIGR: TIGR00631; uvrB; 1.
DR PROSITE: PS50151; UVR; 1.
KW SOS response; Excision nuclease; DNA repair; ATP-binding;
KW Complete proteome.
FT NP_BIND 41 48 ATP (POTENTIAL).
FT DOMAIN 653 688 UVR.
SO SEQUENCE 698 AA; 78158 MW; 7A68B02871502E6F CRC64;

```

```

Alignment Scores:
Pred. No.: 1.45 Length: 698
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 1 Gaps: 0

```

US-09-513-151-3 (1-2041) x UVRB_MYCLE (1-698)

```

QY 95 CTCGGGCGCACGGCGCAATCC 121
    |||||||
DB 40 LeuGlyAlaThrclThrgLysSer 48

```

```

RESULT 4
UVRB_MYCTU STANDARD: PRT: 698 AA.
ID UVRB_MYCTU

```

```

AC 006150;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Excinuclease ABC subunit B.
GN UVRB OR RV1633 OR MT1659 OR MRCY01B2.25.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteriia (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Elgmeyer K., Gas S., Barry C.E. III, Tekela F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton J., Squares R., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RA complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uitterlank T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishel W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RA laboratory strains".
RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRB STIMULATES
CC THE APASE ACTIVITY OF UVRB IN THE PRESENCE OF UV-IRRADIATED
CC DOUBLE-STRANDED DNA. IT ALSO ENHANCES THE ABILITY OF UVRB TO BIND
CC TO UV-IRRADIATED DUPLEX DNA (BY SIMILARITY).
CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVRB, UVRB AND UVRB.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE UVRB FAMILY.
CC -1- SIMILARITY: CONTAINS 1 UVR DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z95554; CAB0886.1;
DR EMBL: AE007030; AAK45939.1; ALT_INIT.
DR HSSP: P56981; 1D9X.
DR TIGR: MT1669;
DR Tuberculist; RV1633;
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR004807; UVRB.
DR InterPro: IPR001943; UVRB/C.
DR Pfam: PF00270; DEAD; 1.
DR Pfam: PF00271; helicase_C; 1.
DR Pfam: PF02151; UVR; 1.
DR Pfam: PF02151; UVR; 1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELIC; 1.
DR TIGR: TIGR00631; uvrB; 1.
DR PROSITE: PS50151; UVR; 1.
KW SOS response; Excision nuclease; DNA repair; ATP-binding;
KW Complete proteome.

```

FT NP_BIND 41 48 ATP (POTENTIAL)
 FT DOMAIN 653 688 UVR.
 SQ SEQUENCE 698 AA; 78070 MW; 9252A0079699CF82 CRC64;
 Alignment Scores:
 Pred. No.: 1.45 Length: 698
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.36% Indels: 0
 DB: 1 Gaps: 0
 US-09-513-151-3 (1-2041) x UVRB_MYCTU (1-698)
 QY 95 CTGCGGGCCAGCGCCGCGCAATCC 121
 DB 40 leuGlyAlaThrGlyThrGlyLysSer 48
 RESULT 5
 RL37_EMENI STANDARD: PRT: 92 AA.
 ID RL37_EMENI STANDARD: PRT: 92 AA.
 AC O9COT1:
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 60S ribosomal protein L37.
 GN RPL37.
 OS *Emicella nidulans* (*Aspergillus nidulans*).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; *Emicella*.
 OX NCBI_TaxID=5072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed:11179686;
 RA Jeong H., Cho G., Han K., Kim J., Min Han D., Jahng K., Chae K.;
 RT "Differential expression of house-keeping genes of *Aspergillus*
 RL *nidulans* during sexual development.";
 RL Gene 262:215-219(2001).
 CC -1- SIMILARITY: BELONGS TO THE L37E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF277382; AK17097.1;
 DR EMBL: AF277381; AK17096.1;
 DR InterPro: IPR001569; Ribosomal_L37E.
 DR Pfam: PF01907; Ribosomal_L37e.1.
 DR ProDom: PD005132; Ribosomal_L37E.1.
 DR PROSITE: PS01077; RIBOSOMAL_L37E; 1.
 KW Ribosomal protein.
 SQ SEQUENCE 92 AA; 10502 MW; 0C924E09573DBFB9 CRC64;
 Alignment Scores:
 Pred. No.: 21.3 Length: 92
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.22% Indels: 0
 DB: 1 Gaps: 0
 US-09-513-151-3 (1-2041) x RL37_EMENI (1-92)
 QY 112 GTGCGCCGTGCGCCGCGAATCAC 89
 DB 85 GYAlaArgGlyProGluAsnHis 92
 RESULT 6
 DIVC_BACSU

ID DIVC_BACSU STANDARD: PRT: 125 AA.
 AC P37471;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cell division protein divC.
 GN DIVC OR DIVA.
 OS *Bacillus subtilis*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=94156852; PubMed=8113187;
 RA Levin P.A., Losick R.;
 RT "Characterization of a cell division gene from *Bacillus subtilis* that
 RL is required for vegetative and sporulation septum formation.";
 RL J. Bacteriol. 176:1451-1459(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=96051385; PubMed=7584024;
 RA Ogasawara N., Nakai S., Yoshikawa H.;
 RT "Systematic sequencing of the 180 kilobase region of the *Bacillus*
 RL *subtilis* chromosome containing the replication origin.";
 RL DNA Res. 1:1-14(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borcherdt S.,
 RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Ertion K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Filiz C., Fujita M., Fujita Y., Funa S., Gallazzi A., Galleron N.,
 RA Gilm S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Bianchini M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kuno M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetle D., Portolillo S., Prescott A.M.,
 RA Prescann E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Stale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpsira P., Tognoni A.,
 RA Taseiro V., Uchiyama S., Vandenbol M., Vannier F., Vasseroletti A.,
 RA Viari A., Wambuit R., Wedler E., Wedler H., Wiltzenegger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RL *subtilis*.";
 RL Nature 390:249-256(1997).
 -1- FUNCTION: REQUIRED FOR VEGETATIVE AND SPORULATION SEPTUM
 FORMATION. REQUIRED FOR THE ACTIVATION OF GENES EXPRESSED UNDER
 THE CONTROL OF THE SPORULATION TRANSCRIPTION FACTORS SIGMA F AND
 SIGMA E.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

CC -----
DR EMBL: L23497; AAB38379.1;
DR EMBL: D26185; BA05297.1;
DR EMBL: 299104; CAB11858.1;
DR Subtilist: BG10125; divIC.
KW Complete proteome.
SQ SEQUENCE 125 AA; 14721 MW; 2ABEB34268B5DC0 CRC64;

Alignment Scores:
Pred. No.: 20.5 Length: 125
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: Gaps: 0

US-09-513-151-3 (1-2041) x DIVC_BACSU (1-125)
OY 1193 AACCACTGAGAGAAAGAGAGA 1216
DB 24 AaGInLeuLysLysArgArgArg 31

RESULT 7
VATD_SULSO STANDARD; PRT; 213 AA.
AC 09UMW9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE V-type ATP synthase subunit D (EC 3.6.3.14) (V-type ATPase subunit D).
GN APPD OR SS00566.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=20165948; PubMed=10701121;
RA Charlebois R.L., Singh R.K., Chan-Weimer C.C.-Y., Allard G., Chow C.,
RA Confalonieri F., Curtis B., Duguet M., Erasuo G., Faguy D.,
RA Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,
RA Kishwana N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,
RA St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F.,
RA Ragan M.A., Sensen C.W.;
RT "Gene content and organization of a 281-kbp contig from the genome of
RT the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";
RL Genome 43:116-136(2000).
RN 121
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=1142726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA De Moers J., Chan-Weimer C.C.-Y., Clausen I.G., Curtis B.A.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -i- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC GRADIENT ACROSS THE MEMBRANE.
CC -i- CATALYTIC ACTIVITY: ATP + H(2O) + H(+)(in) -> ADP + phosphate +
CC H(+)(out).
CC -i- SIMILARITY: BELONGS TO THE V-ATPASE D SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

```

CC -----
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y18930; CAB57734.1;
DR EMBL: AE006687; AAK40881.1;
DR InterPro: IPR002699; ATPsynth_Dsub.
DR Pfam: PF01813; ATP-synth_D; 1.
DR ProDom: PD004122; ATPsynth_Dsub; 1.
DR TIGRFAMs: TIGR00309; V-ATPase_subd; 1.
KW Hydrolase; ATP synthetase; Hydrogen ion transport; Complete proteome.
SQ SEQUENCE 213 AA; 25092 MW; 8943933520729DE8 CRC64;

Alignment Scores:
Pred. No.: 19.2 Length: 213
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: Gaps: 0

US-09-513-151-3 (1-2041) x VATD_SULSO (1-213)
OY 1858 ATTCTACGAGAGAGAGATCA 1881
DB 203 IlleuGlnArgArgArgGluSer 210

RESULT 8
KEY_CHLUM STANDARD; PRT; 216 AA.
AC 09PJUD;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytidylate kinase (EC 2.7.4.14) (CK) (Cytidine monophosphate kinase)
DE (CMP kinase).
GN CMK OR TC0737.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=MOPn / N199;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brumham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson M., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mohn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -i- CATALYTIC ACTIVITY: ATP + (d)CMP -> ADP + (d)CDP.
CC -i- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -i- SIMILARITY: BELONGS TO THE CYTIDYLATE KINASE FAMILY. SUBFAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE002342; AAF39547.1;
DR HSSP: P23863; ICKE.
DR TIGR: TC0737;
DR InterPro: IPR003136; Cytidylate_kin.
DR Pfam: PF02224; Cytidylate_kin; 1.
DR TIGRFAMs: TIGR00017; cmk; 1.
KW Transferase; Kinase; ATP-binding; Complete proteome.
FT NP BIND 7
SQ SEQUENCE 216 AA; 24024 MW; 6013B4CDA7F6C59C CRC64;

Alignment Scores:

```


Pred. No.: 19.1
 Score: 8.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.21%
 DB: 1
 Gaps: 0

US-09-513-151-3 (1-2041) x KCY_CHLTR (1-216)
 QY 107 GGCACCGGCAATCCACGCGCGC 130
 DB 10 GlyThrGlyLysSerThrLeuAla 17

RESULT 9
 KCY_CHLTR
 ID KCY_CHLTR STANDARD: PRT: 216 AA.
 AC 084458:
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytidylate kinase (EC 2.7.4.14) (CK) (Cytidine monophosphate kinase)
 DE (CMP kinase).
 GN CMK OR CT452.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaeae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99000809; PubMed=9784136;
 RA Stephens R.S., Kaiman S., Lammell C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., Davis R.W.;
 RT "Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis";
 RL Science 282:754-759(1998).
 CC -1- CATALYTIC ACTIVITY: ATP + (d)CMP = ADP + (d)CDP.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE CYTIDYLATE KINASE FAMILY. SUBFAMILY 1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AE001319; AAC68052.1; -.
 DR HSP; P23863; ICKE.
 DR InterPro: IPR003136; Cytidylate_kin.
 DR Pfam: PF02224; Cytidylate_kin; 1.
 DR TIGRFAMs: TIGR00017; cmk; 1.
 KM Transferase; Kinase; ATP-binding; Complete proteome.
 FT NP_BIND 7 15
 FT ATP (By Similarity).
 SQ SEQUENCE 216 AA; 24021 MW; A50CB216A9036306 CRC64;

Alignment Scores:
 Pred. No.: 19.1
 Score: 8.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.21%
 DB: 1
 Gaps: 0

US-09-513-151-3 (1-2041) x KCY_CHLTR (1-216)
 QY 107 GGCACCGGCAATCCACGCGCGC 130
 DB 10 GlyThrGlyLysSerThrLeuAla 17

RESULT 10
 KITH_STRCO

ID KITH_STRCO STANDARD: PRT: 216 AA.
 AC 050519:
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Thymidine kinase (EC 2.7.1.21).
 DE TDK OR SC05845 OR SC9810.12.
 GN Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomycetales; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L., Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H., Harper D., Bateman A., Brown A., Chandra G., Chen C.W., Collins M., Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S., Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S., Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S., Seeger K., Saunders D., Sharp S., Squares R., Taylor K., Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J., Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 CC -1- CATALYTIC ACTIVITY: ATP + thymidine = ADP + thymidine 5'-phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: BELONGS TO THE THYMIDINE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AL009204; CAAL5802.1; -.
 DR InterPro: IPR001267; TK_cell.
 DR Pfam: PF00265; TK; 1.
 DR PROSITE: PS00603; TK_CELLULAR_TYPE; FALSE_NEG.
 KM Transferase; Kinase; DNA synthesis; ATP-binding; Complete proteome.
 FT NP_BIND 9 16
 FT ATP (By Similarity).
 SQ SEQUENCE 216 AA; 23368 MW; A04BA1CB8C3B4ADA CRC64;

Alignment Scores:
 Pred. No.: 19.1
 Score: 8.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.21%
 DB: 1
 Gaps: 0

US-09-513-151-3 (1-2041) x KITH_STRCO (1-216)
 QY 113 GGCACCGGCAATCCACGCGCGCGC 136
 DB 14 GlyLysSerThrLeuAlaLeuGln 21

RESULT 11
 YTHB_ERWHE
 ID YTHB_ERWHE STANDARD: PRT: 228 AA.
 AC 047826:
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical protein in tufB 3'region (Fragment).
 OS Erwinia herbicola.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pantoea.

```

OX NCBI_TaxID=549;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 21434;
RA Foor F.;
RL Submitted (May-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE MANNITOL DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U25347; AAA6392.1;
CC InterPro: IPR000669; Mannitol_dh.
CC Pfam: PF01232; Mannitol_dh: 1.
CC PROSITE: PS00974; MANNITOL_DEHYDROGENASE: PARTIAL.
CC Hypothetical protein: Oxidoreductase: NAD.
CC FT NCBIND 48 59 MAD (BY SIMILARITY).
CC NON_TER 228 228
CC SEQUENCE 228 AA; 24588 MW; 0DC5756A85FDF410 CRC64;

Alignment Scores:
Pred. No.: 19 Length: 228
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.22% Indels: 0
DB: 1 Gaps: 0

US-09-513-151-3 (1-2041) x YTDU_ERWHE (1-228)

OY 155 GCAGGAGCCGCTGCGCTAGCTGCA 132
    |||||||
Db 193 ArgArgAlaAlaGlyLeuAlaAla 200

RESULT 12
MTR_A_METMA STANDARD; PRT; 240 AA.
ID MTR_A_METMA
AC OS9640;
DT 15-JUN-1999 (Rel. 38, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Tetrahydromethanopterin S-methyltransferase subunit A (EC 2.1.1.86)
DE (N5-methyltetrahydromethanopterin--coenzyme M methyltransferase
DE subunit A).
GN MTR_A OR M1543.
OS Methanosarcina mazei (Methanosarcina flisla).
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Gol / Gol / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=98218550; PubMed=955648;
RA Lienard T., Gottschalk G.;
RT Cloning, sequencing and expression of the genes encoding the sodium
RT translocating N5-methyltetrahydromethanopterin:coenzyme M
RT methyltransferase of the methylotrophic archaeon Methanosarcina mazei
RT Gol.
RL FEBS Lett. 425:204-208(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Gol / Gol / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppeleier U., Johann A., Hartsch T., Merl R., Schmitz R.A.,
RA Matlener-Artas R., Heme A., Wierer A., Baumeier S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;

```

```

RT "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea."
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC -1- FUNCTION: THIS ENZYME COMPLEX CATALYZES AN INTERMEDIATE STEP IN
CC METHANOGENESIS, THE FORMATION OF METHYL-COENZYME M AND
CC TETRAHYDROMETHANOPTERIN FROM COENZYME M AND N5-METHYL-
CC TETRAHYDROMETHANOPTERIN. THIS IS A ENERGY-CONSERVING, SODIUM-ION
CC TRANSLOCATING STEP.
CC -1- CATALYTIC ACTIVITY: 5-methyl-5,6,7,8-tetrahydromethanopterin + 2-
CC mercaptoethanesulfonate -> 5,6,7,8-tetrahydromethanopterin + 2-
CC (methylthio)ethanesulfonate.
CC -1- COFACTOR: BINDS 5-HYDROXYBENZIMIDAZOL COBAMIDE AS A PROSTHETIC
CC GROUP.
CC -1- PATHWAY: Methanogenesis.
CC -1- SUBUNIT: COMPOSED OF 8 DIFFERENT SUBUNITS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND VESICULAR MEMBRANE-
CC ASSOCIATED (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF042381; AAC38334.1;
CC DR EMBL: AE013388; AAM31239.1;
CC DR TIGRfams: TIGR01111; mtrA: 1.
CC KW Transferase; Methyltransferase; Methanogenesis; Cobalt.
CC FT BINDING 85 85
CC FT 5-HYDROXYBENZIMIDAZOL COBAMIDE COFACTOR
CC (BY SIMILARITY).
CC FT DOMAIN 169 176
CC FT POLY-GLU.
CC FT DOMAIN 232 238
CC FT POLY-LEU.
CC FT CONFLICT 16 16 G -> R (IN REF. 1).
CC FT CONFLICT 71 71 N -> D (IN REF. 1).
CC FT CONFLICT 127 127 E -> G (IN REF. 1).
CC FT CONFLICT 161 161 D -> E (IN REF. 1).
CC SO SEQUENCE 240 AA; 25366 MW; 44C086DD3561E526 CRC64;

Alignment Scores:
Pred. No.: 18.9 Length: 240
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.22% Indels: 0
DB: 1 Gaps: 0

US-09-513-151-3 (1-2041) x MTR_A_METMA (1-240)

OY 1034 GCAGGATTTCAGAGCAGCTTCAA 101
    |||||||
Db 123 AlaArgPheGlnGlnValGln 130

RESULT 13
NUDC_HAEIN STANDARD; PRT; 264 AA.
ID NUDC_HAEIN
AC P44710;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADH pyrophosphatase (EC 3.6.1.-).
GN NUDC OR H10432
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

```

RA Mckenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodok A., Kelley J.M.,
 RA Weiman J.F., Phillips C.A., Spriggs T., Hebbloom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Geophagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd";
 RL Science 269:496-512(1995).
 CC -I- CATALYTIC ACTIVITY: NADH + H(2)O = AMP + NMNH.
 CC -I- COPACITOR: REQUIRES DIVALENT IONS: MANGANESE OR MAGNESIUM (BY
 CC SIMILARITY).
 CC -I- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO THE NUDIX HYDROLASE FAMILY. NUDC SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32726; AAC22091.1; .
 CC TIGR: H10432; .
 DR InterPro: IPR000086; NUDIX_hydrolase.
 DR Pfam: PF00293; NUDIX.1.
 DR PRINTS: PR00302; NUDIXFAMILY.
 DR PROSITE: PS00893; NUDIX.1.
 KW Hydrolyase; NAD; Magnesium; Manganese; Complete proteome.
 FT DOMAIN 163 184 NUDIX BOX.
 SQ SEQUENCE 264 AA; 30163 MW; 1720219015B9922E CRC64;
 Alignment Scores:
 Pred. No.: 18.6 Length: 264
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.21% Indels: 0
 DB: Gaps: 0
 US-09-513-151-3 (1-2041) x NUDC_HAELIN (1-264)
 Oy 581 GTTTTGAGAAACAGATCTCT 604
 Db 179 ValPheGluGluThrclyIleSer 186
 RESULT 14
 ID PLCD_HUMAN STANDARD: PRT; 378 AA.
 AC Q9NRZ5:
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase delta (EC 2.3.1.51) (1-
 DE AGP acyltransferase 4) (1-AGPAT 4) (lysophosphatidic acid
 DE acyltransferase-delta) (LPAAT-delta) (1-acylglycerol-3-phosphate O-
 DE acyltransferase 4).
 GN AGPAT4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCB1_Taxid:9606;
 RN NCB1_Taxid:9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Leung D.W.;
 RT "Structure and functions of lysophosphatidic acid acyltransferases";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN (12)
 RC SEQUENCE FROM N.A.
 RC TISSUE-Spleen;
 RA Strausberg R.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY
 CC SIMILARITY).
 CC -I- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.
 CC -I- PATHWAY: De novo phospholipid biosynthesis; second step.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -I- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
 CC ACYLTRANSFERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF156776; AAF80338.1; .
 DR EMBL: BC020203; AAH20209.1; .
 DR InterPro: IPR002123; Acyltransferase.
 DR Pfam: PF01553; Acyltransferase.1.
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;
 FT TRANSMEM 11 31 POTENTIAL.
 FT TRANSMEM 125 145 POTENTIAL.
 FT TRANSMEM 307 327 POTENTIAL.
 FT TRANSMEM 338 358 POTENTIAL.
 SQ SEQUENCE 378 AA; 44021 MW; 3EFCD13D196F8CDE5 CRC64;
 Alignment Scores:
 Pred. No.: 17.8 Length: 378
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.21% Indels: 0
 DB: Gaps: 0
 US-09-513-151-3 (1-2041) x PLCD_HUMAN (1-378)
 Oy 669 AGTCTCTAACCTGCATCCTTT 692
 Db 334 SerSerLeuThrLeuLasePhe 341
 RESULT 15
 ID YP28_CAEEL STANDARD: PRT; 417 AA.
 AC P98061;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Hypothetical zinc metalloproteinase F42A10.8 precursor (EC 3.4.24.-).
 GN F42A10.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NC NCB1_Taxid:6239;
 RN NCB1_Taxid:6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA STRAIN-Bristol N2;
 RL Latreille P.;
 RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
 CC -I- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 CUB DOMAIN.
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@sib.ch).

CC -----
 DR EMBL: U10414; AAA19079.1; -
 DR HSSP: P28825; 11AF.
 DR MEROPS: M12.0PA; -
 DR WormRep: F42A10.8; CE01299.
 DR InterPro: IPR001506; Astacin.
 DR InterPro: IPR000859; CUB-domain.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR00030; Zn_MTPeptide.
 DR Pfam: PF00431; CUB; 1.
 DR Pfam: PF01400; Astacin; 1.
 DR SMART: SM00042; CUB; 1.
 DR SMART: SM00235; Zmc; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR PROSITE: PS0180; CUB; 1.
 DR PROSITE: PS00022; EGF-1; 1.
 DR PROSITE: PS0186; EGF-2; FALSE_NEG.
 KW Hypothetical protein; Repeat; Hydrolase; Metalloprotease;
 KW EGF-like domain; Zinc; Signal.
 FT SIGNAL 1 14
 FT CHAIN 15 417
 FT METAL 134 134
 FT ACT_SITE 135 135
 FT METAL 138 138
 FT METAL 144 144
 FT DOMAIN 244 274
 FT DISULFID 248 259
 FT DISULFID 251 262
 FT DISULFID 264 273
 FT CARBOHYD 76 76
 FT CARBOHYD 237 237
 FT CARBOHYD 314 314
 SQ SEQUENCE 417 AA: 46790 MW: A0DE11E2B628C98D CRC64;
 Alignment Scores:
 Pred. No.: 17.6
 Score: 8.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.22%
 DB: 1
 Length: 417
 Matches: 8
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-09-513-151-3 (1-2041) x YP28_CAEEL (1-417)

OY 486 CCATCCTCTTTCAGCTCACT 463
 |||||
 Db 305 ProSerSerSerSerSerThr 312

Search completed: April 21, 2003, 19:15:30
 Job time : 29.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 19:12:25 : Search time 44 seconds

(without alignments)
8918.653 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 661

Sequence: 1 CTGCCATAGAGCGGCTCCG.....TTTACAGAGAAAAAAA 2041

Scoring table:

OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283224 segs, 96134422 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565536

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame-n2p.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US095151/runal_15042003_141247_27235/app_query.fasta_1.2183
-DB=pir_73 -OFMT=fasta -SUFFIX=olin2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=olin2p -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=15 -MODE=LOCAL -OUFFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0951513151.qcgn_1_1_52_etunal_15042003_141247_27235 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database:

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	9	1.4	196	2	S47088	psad protein - Chd
2	9	1.4	382	2	S54702	twitching motility
3	9	1.4	428	2	S67176	tRNA isopentenyltr
4	9	1.4	434	2	T38664	tRNA isopentenyltr
5	9	1.4	698	2	G70559	probable uvrB prot
6	9	1.4	698	2	E87082	exonuclease ABC s
7	8	1.2	112	2	D81432	hypothetical prote
8	8	1.2	121	2	C71104	hypothetical prote
9	8	1.2	125	2	B53380	cell division init
10	8	1.2	162	2	G75497	hypothetical prote
11	8	1.2	213	2	B90203	ATP synthase subun
12	8	1.2	216	2	T35881	thymidine kinase (
13	8	1.2	216	2	G71512	probable cmp kinas
14	8	1.2	216	2	E81670	cytidylate kinase

15	8	1.2	246	2	F72464	hypothetical prote
16	8	1.2	263	2	G84083	ABC transporter (A
17	8	1.2	264	1	F64152	hypothetical prote
18	8	1.2	277	2	T05658	hypothetical prote
19	8	1.2	301	2	S63661	hypothetical prote
20	8	1.2	320	2	G96714	hypothetical prote
21	8	1.2	330	2	T48100	tRNA isopentenyl t
22	8	1.2	357	2	F96708	hypothetical prote
23	8	1.2	372	2	S17715	transcription acti
24	8	1.2	374	2	B82572	P-protein XF2325 (
25	8	1.2	417	2	T30863	hypothetical prote
26	8	1.2	433	2	T05693	hypothetical prote
27	8	1.2	477	2	C59098	phosphomannomutase
28	8	1.2	483	2	A12868	phosphomannomutase
29	8	1.2	498	2	B70759	hypothetical prote
30	8	1.2	507	2	C97645	phosphomannomutase
31	8	1.2	507	2	S43877	trwB protein - Esc
32	8	1.2	536	2	G85488	probable transport
33	8	1.2	536	2	G90637	probable transport
34	8	1.2	580	2	S34340	NPL4 protein - yea
35	8	1.2	656	2	A64208	exonuclease ABC c
36	8	1.2	657	2	S73946	exonuclease ABC c
37	8	1.2	658	2	AC1761	exonuclease ABC c
38	8	1.2	658	2	A11385	exonuclease ABC c
39	8	1.2	660	2	C84099	exonuclease ABC c
40	8	1.2	661	2	G69729	exonuclease ABC c
41	8	1.2	663	2	G89848	exonuclease ABC su
42	8	1.2	665	2	A11947	exonuclease ABC c
43	8	1.2	666	2	C82932	exonuclease ABC s
44	8	1.2	669	2	S74391	exonuclease ABC c
45	8	1.2	674	2	T01309	probable serine/th

ALIGNMENTS

RESULT 1

S47088

psad protein - Chlamydomonas reinhardtii

C:Species: Chlamydomonas reinhardtii

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 26-Aug-1999

C:Accession: S47088

R:Farrah, J.A.; Frank, G.; Zuber, H.; Rochaix, J.D.

Submitted to the EMBL Data Library, June 1994

A:Description: Cloning and sequencing of a cDNA encoding the photosystem I Psad s

A:Reference number: S47088

A:Accession: S47088

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-196 <FAR>

A:Cross-references: EMBL:X79674; NID:9498823; PIDN:CAA56122.1; PID:9498824

C:Superfamily: photosystem I chain II

Alignment Scores:

Pred. No.: 3.84
Score: 9.00
Percent Similarity: 100.00%
Best local Similarity: 100.00%
Query Match: 1.36%
DB: 2
Gaps: 0

US-09-513-151-3 (1-2041) x S47088 (1-196)

OY 130 GTTGCAGCTAGCCAGCGCTCCGCG 156

|||||

Db 21 ValAlaAlaArgProAlaAlaArgArg 29

RESULT 2

S54702

twitching motility protein PilU PA0396 [Imported] - Pseudomonas aeruginosa (strain

C:Species: Pseudomonas aeruginosa

C:Date: 06-Sep-1996 #sequence_revision 13-Mar-1997 #text_change 24-Oct-2000

C:Accession: S54702; B83595

R:Whitchurch, C.B.; Mattick, J.S.

MOL: Microbiol. 13, 1079-1091, 1994
 A:Title: Characterization of a gene, pilU, required for twitching motility but not phage
 A:Reference number: S54702; MUID:95157252; PMID:7854122
 A:Accession: S54702
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-382 <WHI>
 A:Cross-references: EMBL:L27667; NID:q443685; PIDN:AAA25965.1; PID:q443686
 R:Stover, C.K.; Pham, X.Q.; Ewin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
 .; Loy, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: B83595
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-382 <STO>
 A:Cross-references: GB:AE004477; GB:AE004091; NID:g9946248; PIDN:ANG03785.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: pilU; PA0396
 C:Superfamily: twitching motility protein pilT

Alignment Scores:
 Pred. No.: 3.52 Length: 382
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.36% Indels: 0
 DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x S54702 (1-382)

QY 98 GGGGCCACGGCCCGGCAATCCACG 124
 Db 130 GYAlAThrglyThrglyLysSerThr 138

RESULT 3
 S67176
 tRNA isopentenyltransferase (EC 2.5.1.8) - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein 05447w; protein YOR274w
 C:Species: Saccharomyces cerevisiae
 C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 21-Jul-2000
 C:Accession: S67176; A26717; S72045
 R:Cherett, G.; Sor, F.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67169
 A:Accession: S67176
 A:Molecule type: DNA
 A:Residues: 1-428 <CHE>
 A:Cross-references: EMBL:Z75182; NID:q1420613; PID:e252418; PID:q1420614; MIPS:YOR274w
 A:Experimental source: strain S288C
 R:Naarajan, D.; Dhanich, M.E.; Martin, N.C.; Hopper, A.K.
 Mol. Cell. Biol. 7, 185-191, 1987
 A:Title: DNA sequence and transcript mapping of MOD5: features of the 5' region which su
 A:Reference number: A26717; MUID:87172703; PMID:3031457
 A:Accession: A26717
 A:Molecule type: DNA
 A:Residues: 1-374, R, 376-428 <NAU>
 A:Cross-references: EMBL:M15991
 R:Cherett, G.; Bernard, A.; Sor, F.
 Yeast 12, 1059-1064, 1996
 A:Title: DNA sequence analysis of the VPB1-SNF2 region on chromosome XV of Saccharomyces
 A:Reference number: S72039; MUID:97051594; PMID:8896271
 A:Accession: S72045
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-428 <CHW>
 A:Cross-references: EMBL:X89633; NID:q1279694; PIDN:CAA61780.1; PID:q1419759
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995
 C:Genetics:
 A:Gene: SGD:MOD5

A:Cross-references: MIPS:YOR274w; SGD:S0005800
 A:Map position: 15R
 C:Keywords: transferase

Alignment Scores:
 Pred. No.: 3.46 Length: 428
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.36% Indels: 0
 DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x S67176 (1-428)

QY 350 AAAATTCCTATTGTTGGAGGACG 376
 Db 104 LysileProileValValGlyGlyThr 112

RESULT 4
 T38664
 trna isopentenyltransferase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T38664
 R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: Z21804
 A:Accession: T38664
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-434 <MUR>
 A:Cross-references: EMBL:AL109739; NID:e1534774; PIDN:CAB52278.1; GSPDB:GN00066; ;
 A:Experimental source: strain 972h; cosmid c343
 C:Genetics:
 A:Gene: SPDB:SPAC343.15
 A:Map position: 1

Alignment Scores:
 Pred. No.: 3.46 Length: 434
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.36% Indels: 0
 DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x T38664 (1-434)

QY 350 AAAATTCCTATTGTTGGAGGACG 376
 Db 93 LysileProileValValGlyGlyThr 101

RESULT 5
 G70559
 probable uvrB protein - Mycobacterium tuberculosis (strain H37Rv)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 02-Feb-2001
 C:Accession: G70559
 R:Cherett, G.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gor
 rajandream, M.A.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holto
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete g
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: G70559
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-698 <COL>
 A:Cross-references: GB:Z95554; GB:AL123456; NID:g361771; PIDN:CAB08886.1; PID:g21
 A:Experimental source: strain H37Rv
 C:Genetics:
 A:Gene: uvrB
 C:Superfamily: excinuclease ABC chain B

C:Keywords: ATP; nucleotide binding; P-loop
 F:41-48/Region: nucleotide-binding motif A (P-loop)
 F:335-340/Region: nucleotide-binding motif B
 F:339-342/Region: DEXH motif

Alignment Scores:
 Pred. No.: 3.24 Length: 698
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.36% Indels: 0
 DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x G70559 (1-698)

OY 95 CTCGGGGCCACGGCCACCGCAATCC 121
 |||||
 Db 40 LeuGlyAlaThrGlyThrGlyLysSer 48

RESULT 6

E87082
 excinuclease ABC subunit B [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae
 C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001

C:Accession: E87082

R:Coile, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
 R.; Davies, R.M.; Devlin, K.; Dutfoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
 eam, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001

A:Authors: Rutter, S.; Seeger, K.; Simmonds, M.; Skelton, J.; Squares, R.; Sc
 A:Title: Massive gene decay in the leprosy bacillus.

A:Reference number: A86909; MUID:21128732; PMID:11234002

A:Accession: E87082

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-698 <STO>

A:Cross-references: GB:AL450380; NID:g13093274; PIDN:CAC31768.1; GSPDB:GN00147

C:Genetics:

A:Gene: uvrB

C:Superfamily: excinuclease ABC chain B

Alignment Scores:
 Pred. No.: 3.24 Length: 698
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.36% Indels: 0
 DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x E87082 (1-698)

OY 95 CTCGGGGCCACGGCCACCGCAATCC 121
 |||||
 Db 40 LeuGlyAlaThrGlyThrGlyLysSer 48

RESULT 7

D81432
 hypothetical protein Cj0148c [imported] - Campylobacter jejuni (strain NCTC 11168)

C:Species: Campylobacter jejuni

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002

C:Accession: D81432

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chiller
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyl

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: D81432

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-112 <PAR>

A:Cross-references: GB:AL139074; GB:AL111168; NID:g6967505; PIDN:CAB72632.1; PID:g696764

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj0148c
 C:Superfamily: Campylobacter jejuni hypothetical protein Cj0148c

Alignment Scores:
 Pred. No.: 46.6 Length: 112
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.21% Indels: 0
 DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x D81432 (1-112)

OY 573 GCTGCAAGTTTGAAGAAACAG 596
 |||||
 Db 16 AlecysLysPheLeuLysLysGln 23

RESULT 8

C71104
 hypothetical protein PH1094 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000

C:Accession: C71104

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halykawa, Y.; Hino, Y.; Yamamoto, S.
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kishida, N.;

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermop

A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: C71104

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-121 <KAW>

A:Cross-references: GB:AP000004; NID:g3226131; PIDN:BAA30193.1; PID:g3257510

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by c

C:Genetics:

A:Gene: PH1094

C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1094

Alignment Scores:
 Pred. No.: 46.2 Length: 121
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.22% Indels: 0
 DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x C71104 (1-121)

OY 1111 ACTCTCTGTTCTCAGCTTCATT 1088
 |||||
 Db 6 ThierLeuValLeuSerPheIle 13

RESULT 9

B53380
 cell division initiation protein divIC - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 20-Jun-2000

C:Accession: B53380; S66092; A69616

R:Levin, P.A.; Losick, R.

J. Bacteriol. 176, 1451-1459, 1994

A:Title: Characterization of a cell division gene from Bacillus subtilis that is

A:Reference number: A53380; MUID:94156852; PMID:8113187

A:Accession: B53380

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-125 <LEUV>

A:Cross-references: GB:L23497; NID:g469178; PIDN:AAB38379.1; PID:g385177

R:Ogasawara, N.; Nakai, S.; Yoshikawa, H.

DNA Res. 1, 1-14, 1994

A:Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis

A:Reference number: S65967; MUID:96051385; PMID:7584024

A:Accession: S66092

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-125 <OGA>
A:Cross-references: EMBL:D6185; NID:g467326; PIDN:BA05397.1; PID:g467451
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bartet
C.; Biron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabel, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogihara, A.; Oudeg, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schreier, R.; Scorfone, F.; Sekiguchi, J.; Sekowska, A.; Ser
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Walters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Dancho, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580, MUID:98044033; PMID:9384377
A:Accession: A69616
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-125 <RUN>
A:Cross-references: GB:Z99104; GB:AL009126; NID:g2632267; PIDN:CAB11838.1; PID:g2632329
A:Experimental source: strain 168
C:Genetics:
A:Gene: divIC
A:Start codon: TTG
C:Superfamily: *Bacillus subtilis* cell division initiation protein divIC

Alignment Scores:
Pred. No.: 46 Length: 125
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x B53380 (1-125)
QY 1193 AACCAACTGAGAGAGAGAGACA 1216
|||||
DB 24 AaNgInLeuLysLysArGArGArG 31

RESULT 10
C75497
Hypothetical protein - *Deinococcus radiodurans* (strain R1)
C:Species: *Deinococcus radiodurans*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: C75497
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: C75497
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-162 <WHI>
A:Cross-references: GB:AE001919; GB:AE000513; NID:g6458307; PIDN:AAF10196.1; PID:g645831
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0609
A:Map position: 1

Alignment Scores:
Pred. No.: 44.4 Length: 162
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0

DB: 2 Gaps: 0
US-09-513-151-3 (1-2041) x C75497 (1-162)
QY 107 GCGACCGCAATTCACCGCGCG 130
|||||
DB 84 GlyThrGlyLysSerThrLeuAla 91

RESULT 11
B90203
ATP synthase subunit D (atpD) [imported] - *Sulfolobus solfataricus*
C:Species: *Sulfolobus solfataricus*
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 02-Aug-2002
C:Accession: B90203
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Re
arrett, R.A.; Ragan, M.A.; Sersen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: *Sulfolobus solfataricus* complete genome.
A:Reference number: A99139
A:Accession: B90203
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <KUR>
A:Cross-references: GB:AE006641; NID:g13813731; PIDN:AAK40881.1; GSPDB:GN00155
C:Genetics:
A:Gene: atpD
C:Superfamily: H⁺-transporting ATPase chain D

Alignment Scores:
Pred. No.: 42.8 Length: 213
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x B90203 (1-213)
QY 1858 ATCTACAGAGAGAGGAATCA 1881
|||||
DB 203 IleLeuGlnArGArGArGLeuSer 210

RESULT 12
T35881
thymidine kinase (EC 2.7.1.21) [similarity] - *Streptomyces coelicolor*
C:Species: *Streptomyces coelicolor*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jun-2002
C:Accession: T35881
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Raghindram, M.A.
submitted to the EMBL Data Library, November 1997
A:Reference number: 221592
A:Accession: T35881
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-216 <OLI>
A:Cross-references: EMBL:AL009204; PIDN:CAA15802.1; GSPDB:GN00070; SCOEDB:SC9B10.1
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC9B10.12
C:Superfamily: *Escherichia coli* thymidine kinase
C:Keywords: phosphotransferase

Alignment Scores:
Pred. No.: 42.7 Length: 216
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x T35881 (1-216)

OY 113 GGCAATCCAGCTGGCTTCAG 136
 |||||||||||||||||||
 DB 14 GilyserthrleuAlaLeuGln 21

RESULT 13

G71512

Probable cmp kinase - Chlamydia trachomatis (serotype D, strain UW3/Cx)

C:Species: Chlamydia trachomatis

C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999

C:Accession: G71512

R:Stephens, R.S.; Kaiman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,

Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac

A:Reference number: A71570; MUID:99000809; PMID:9784136

A:Accession: G71512

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-216 <ARN>

A:Cross-references: GB:AE001319; GB:AE001273; NID:q3328881; PIDN:ACG68052.1; PID:q332888

A:Experimental source: serotype D, strain UW-3/Cx

C:Genetics:

A:Gene: cmk

C:Superfamily: cytidylate kinase cmk

Alignment Scores:

Pred. No.:	42.7	Length:	216
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	2	Gaps:	0

US-09-513-151-3 (1-2041) x G71512 (1-216)

OY 107 GGCACCGCAATCCAGCTGGCG 130

|||||

DB 10 GilythrGilyllysserthrleuAla 17

RESULT 14

E81670

cytidylate kinase TC0737 [imported] - Chlamydia muridarum (strain Nigg)

C:Species: Chlamydia muridarum, Chlamydia trachomatis MOPn

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000

C:Accession: E81670

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,

C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MOPn and Chlamydia pneumoniae AR39.

A:Reference number: A81500; MUID:20150255; PMID:10684935

A:Accession: E81670

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-216 <TET>

A:Cross-references: GB:AE002342; GB:AE002160; NID:g7190763; PIDN:AAF39547.1; PID:g719076

A:Experimental source: strain Nigg (MOPn)

C:Genetics:

A:Gene: TC0737

C:Superfamily: cytidylate kinase cmk

Alignment Scores:

Pred. NO.:	42.7	Length:	216
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	2	Gaps:	0

US-09-513-151-3 (1-2041) x E81670 (1-216)

OY 107 GGCACCGCAATCCAGCTGGCG 130

|||||

DB 10 GilythrGilyllysserthrleuAla 17

RESULT 15

F72464

hypotheical protein APE2361 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: F72464

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon,

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: F72464

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-246 <KAW>

A:Cross-references: DDBJ:AP000064; NID:q5105945; PIDN:BAAB1374.1; PID:q5106063

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2361

C:Superfamily: peroxidase

Alignment Scores:

Pred. No.:	42	Length:	246
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	2	Gaps:	0

US-09-513-151-3 (1-2041) x F72464 (1-246)

OY 752 CTTGCTGCTGGCTCTTGAGAGGA 775

|||||

DB 193 leuAlaAlaGilyleuLeuGln 200

Search completed: April 21, 2003, 19:21:02
 Job time : 51 secs

BASE COUNT	224 a	145 c	165 g	177 t
ORIGIN				
Query Match	29.1%;	Score 594.8;	DB 12;	Length 711;

Search completed: April 21, 2003, 18:21:11
Job time : 3858.56 secs

Seq primer: M13 FORWARD
POLYA-Tes.

FEATURES
source

Location/Qualifiers
1..624
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5845067"
/clone_11b="NCI_CGAP_E11"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Left Pelvis; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI_CGAP_E11 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma. The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (d)T18 tail. The sequence tag for this library is
ACACTGCAC.
TAG_L1B-U1-H-E11
TAG_TISSUE=chondrosarcoma
TAG_SEO=ACACTGCAC"

BASE COUNT 203 a 128 c 97 g 195 t 1 others
ORIGIN

Query Match 29.7%; Score 606.8; DB 14; Length 624;
Best Local Similarity 99.4%; Pred. No. 2e-111;
Matches 619; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1419 AGGAGTGCATGCGGGAATCTCTGCATACAGAAAAGCTCCACCATTCTTTGATGT 1478
DB 624 AGGAGTGCATGCGGGAATCTCTGCATACAGAAAAGCTCCACCATTCTTTGATGT 565
QY 1479 GGTTTAAAGTCTCAGCTTCTCTAATATAGAAAACAGAGTCTTGACAGCTCTGTG 1538
DB 564 GGTTTAAAGTCTCAGCTTCTCTAATATAGAAAACAGAGTCTTGACAGCTCTGTG 505
QY 1539 GGTGATGTCCTGGAATATGATGATGTCAGGAAAGCAATTTTTTTCTTTGAACCTTAA 1598
DB 504 GGTGATGTCCTGGAATATGATGATGTCAGGAAAGCAATTTTTTTCTTTGAACCTTAA 445
QY 1599 AGGTCTATTATTAAGAGAGACAGATCCACATTTTATCATAGATCTCTCTTGT 1658
DB 444 AGGTCTATTATTAAGAGAGACAGATCCACATTTTATCATAGATCTCTCTTGT 385
QY 1659 GGTGATACAGAGATGATGATCCCTTTAAAGAGTTTATGTCCTGACAGCTGGCT 1718
DB 384 GGTGATACAGAGATGATGATCCCTTTAAAGAGTTTATGTCCTGACAGCTGGCT 326
QY 1719 AAAATATCTAATTCAGATGCTTTTGTAGATGACTGAAGATTTGTAGCCACATATT 1778
DB 325 AAAATATCTAATTCAGATGCTTTTGTAGATGACTGAAGATTTGTAGCCACATATT 266
QY 1779 GGGAGTCTAGATTTGAGTGAATGGCAGAAAGGCCATCTCCATTGAGATGTTAAGTG 1838
DB 265 GGGAGTCTAGATTTGAGTGAATGGCAGAAAGGCCATCTCCATTGAGATGTTAAGTG 206
QY 1839 AACCAAACTAGTTCGGAATCTACAGAGAGAGGAGGAATACAGCTGAGGAAGCTGA 1898
DB 205 AACCAAACTAGTTCGGAATCTACAGAGAGAGGAGGAATACAGCTGAGGAAGCTGA 146
QY 1899 CATAGAGTCTGAGAGCAAGACTTTGAATTTGGAGCTGCATGCTGATTTAT 1958
DB 145 CATAGAGTCTGAGAGCAAGACTTTGAATTTGGAGCTGCATGCTGATTTAT 86
QY 1959 CACTGCTGCTTTCTATTGATGATCAATCTAATTTTATGAACTTTAAATAAGAA 2018

DB 85 CACTGCTGCTTTCTATTGATGATCAANCTAATTTTATGAACTTTAAATAAGAA 26
QY 2019 AAATTACAGAAAAA 2041
DB 25 AAATTACAGAAAAA 3

RESULT 14
AV759288 639 bp mRNA linear EST 19-OCT-2000
LOCUS AV759288 MDS Homo sapiens cDNA clone MDSBLH01.5, mRNA sequence.
DEFINITION AV759288
ACCESSION AV759288
VERSION AV759288.1 GI:10917136
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
1 (bases 1 to 639)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

TITLE
JOURNAL
COMMENT
Unpublished (2000)
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzge@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..639
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="MDSBLH01"
/clone_11b="MDS"
/tissue_type="bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/note="Vector: pT73-Pac; Site_1: SfiI; Site_2: SfiI"

BASE COUNT 166 a 158 c 175 g 139 t 1 others
ORIGIN

Query Match 29.4%; Score 599.8; DB 10; Length 639;
Best Local Similarity 98.4%; Pred. No. 5.1e-110;
Matches 628; Conservative 0; Mismatches 3; Indels 7; Gaps 2;

QY 1 CTGCCATTAAGATGGCGTCCGTCGTCGACAGAGAGTTCTGTGGAGAGTGGCTCAG 60
DB 5 CTGCCATTAAGATGGCGTCCGTCGTCGACAGAGAGTTCTGTGGAGAGTGGCTCAG 64
QY 61 GGGCTGCAACGAGCCCTACCTCTTGTAGTATCTCGGGGCCACGCGCAACCGCAATC 120
DB 65 GGGCTGCAACGAGCCCTACCTCTTGTAGTATCTCGGGGCCACGCGCAACCGCAATC 124
QY 121 CACGCTGGCGTTGACAGTACGCGGCTCGCGGCTGAGATGTACGCGCTACTTCAT 180
DB 125 CACGCTGGCGTTGACAGTACGCGGCTCGCGGCTGAGATGTACGCGCTACTTCAT 184
QY 181 GCAAGTCTAAGAGGCTGACATCAATCAACCAAGGTTCTGCCAAGAGAGAGAAAT 240
DB 185 GCAAGTCTAAGAGGCTGACATCAATCAACCAAGGTTCTGCCAAGAGAGAGAAAT 244
QY 241 CTGCCGACCAACATGATGAGCTTTGTGATCTCTTGTGACCAATTAACAGTGGTGA 300
DB 245 CTGCCGACCAACATGATGAGCTTTGTGATCTCTTGTGACCAATTAACAGTGGTGA 304
QY 301 CTTCAGAAATTAAGCAACTGCTCTGATTAAGATATATTTGGCCGAGACAAAATCTCAT 360
DB 305 CTTCAGAAATTAAGCAACTGCTCTGATTAAGATATATTTGGCCGAGACAAAATCTCAT 364

Db 482 TTGTACTTCACAACCGCTTAAGCCAGGTGGAGCCAGAACATGGCTGCCAATGTCATCA 541

QY 548 CANGAAGAAAGCGAAAGTGGCGGAGCTTGGCAAGTGGTGAAGAACAGAAATCTGCAT 607

Db 542 CAGTAAGAAAGCGAAAGTGGCGGAGCTTGGCAAGTGGTGAAGAACAGAAATCTGCAT 601

QY 608 AGTGAATTTCTCCATGCTCAACATACGGAAGAGTGGTGGTCCCTTGGAGGCTCTGTG 667

Db 602 AGTGAATTTCTCCATGCTCAACATACGGAAGAGTGGTGGTCCCTTGGAGGCTCTGTG 661

QY 668 AAGTCTCTAACCCTTGCATCCCTTGGCTTCATGCTGACGAGCAGTGTAGATGAGCG 727

Db 662 AGTCTCTAACCCTTGCATCCCTTGGCTTCATGCTGACGAGCAGTGTAGATGAGCG 718

QY 728 TTGGATTAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 763

Db 719 TGGCATAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 754

RESULT 12

LOCUS B0034660 772 bp mRNA linear EST 24-JAN-2001

DEFINITION 602300022F1 NIH_MGC.87 Homo sapiens cDNA clone IMAGE:4394260 5', mRNA sequence.

ACCESSION B0034660

VERSION B0034660.1 GI:12428227

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgc.nci.nih.gov/

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgapds-remail.nih.gov

Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: http://image.lnll.gov

Plate: LMA10089 row: 1 column: 05

High quality sequence stop: 692.

FEATURES

source

1. 772

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4394260"

/clone_id="NIH_MGC.87"

/tissue_type="mammary adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: breast; Vector: pCMV-SPORT6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 224 a 155 c 178 g 214 t 1 others

ORIGIN

Query Match 30.88; Score 628; DB 12; Length 772;

Best Local Similarity 95.18; Pred. No. 1,le-115;

Matches 735; Conservative 0; Mismatches 26; Indels 12; Gaps 8;

QY 1001 GAGCTGTGCTGAACTGCTGGAATGCTGCAAGTTTCATCCAGGGCCACAAGCT 1060

Db 2 GAGCTGTGCTGAACTGCTGGAATGCTGCAAGTTTCATCCAGGGCCACAAGCT 60

QY 1061 ACAGCCACTCCATTAAGATGCCATACATGAAAGCTGAGAACAGAGAAGTTATCACCTG 1120

Db 61 ACAGCCACTCCATTAAGATGCCATACATGAAAGCTGAGAACAGAGAAGTTATCACCTG 120

QY 1121 TGTGACCTCTGTGATGCAATCATCATTTGGGATCGGAATGGGACGCACATTAATGCC 1180

Db 121 TGTGACCTCTGTGATGCAATCATCATTTGGGATCGGAATGGGACGCACATTAATGCC 180

QY 1181 AAATCCACTTGAACCACTGAAGAAAAGAAAGATTTGGACTGACATGCTGTCAACACC 1240

Db 181 AAATCCACTTGAACCACTGAAGAAAAGAAAGATTTGGACTGACATGCTGTCAACACC 240

QY 1241 ATGAAGATGAGATGTTTCCCACTATAACAAAGAACTTAAGGAAGGATGCCCA 1300

Db 241 ATGAAGATGAGATGTTTCCCACTATAACAAAGAACTTAAGGAAGGATGCCCA 300

QY 1301 GGCAGATATGATCAAGAGCTGGAATGAGCGTTTAAAGACATGTCAGTGGCTTTGGA 1360

Db 301 -GGCAGATATGATCAAGAGCTGGAATGAGCGTTTAAAGACATGTCAGTGGCTTTGGA 359

QY 1361 AAGTGTGGGAGATCCAGTTCAAGAGGAGGAGGTATGTTGTCTCCAGTCTGGCAAG 1420

Db 360 AAGTGTGGGAGATCCAGTTCAAGAGGAGGAGGTATGTTGTCTCCAGTCTGGCAAG 419

QY 1421 GAGTGTATGCGGATTTCTGATAGCAGACAAAAGCTCCACATTTCTTTGATGTG 1480

Db 420 GAGTGTATGCGGATTTCTGATAGCAGACAAAAGCTCCACATTTCTTTGATGTG 479

QY 1481 TTTTAAAGTCTCAAGTCTCTATATATGAAGAACAGAGCTTGTGACGCTCTGTGTCG 1540

Db 480 TTTTAAAGTCTCAAGTCTCTATATATGAAGAACAGAGCTTGTGACGCTCTGTGTCG 539

QY 1541 TGATGTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1599

Db 540 TGATGTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 599

QY 1600 GGTCTCTATATTAAGACA -GCACAGATTCACCA -TTTTATATCATAGATGCTCTG 1657

Db 600 GGTCTCTATATTAAGACA -GCACAGATTCACCA -TTTTATATCATAGATGCTCTG 659

QY 1658 TGGTGAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1717

Db 660 TGGTGAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 715

QY 1718 TAAATATATATATTTCCAGATGCTTTTGTAGATGACTGAAGATTTGTGAGC 1770

Db 716 CTAAT -ATCTAATTCAGATGCTTTGTGTA -ATGACTGAGTGTGTGTGAC 765

RESULT 13

LOCUS B0003256/c 624 bp mRNA linear EST 26-MAR-2002

DEFINITION UI-H-E11-ayx-n-12-0-UI.s1 NCI CGAP_E11 Homo sapiens cDNA clone. IMAGE:5845067 3', mRNA sequence.

ACCESSION B0003256

VERSION B0003256.1 GI:19728156

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

COMMENT Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapds-remail.nih.gov

Tissue Procurement: Dr. Jose Mercuende

CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: http://image.lnll.gov

The following repetitive elements were found in this cDNA sequence: 1-36, >POLY_A#simple_repeat (matched complement)

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
(bases 1 to 957)	NIH-MGC	http://mgc.ncl.nih.gov/	National Institutes of Health, Mammalian Gene Collection (MGC)	
Unpublished (1999)	Contact: Robert Strausberg, Ph.D.	Email: cgaps-remail.nih.gov	Tissue Procurement: ARCC	
CDNA Library Preparation: Ling Hong/Rubin Laboratory	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)	DNA Sequencing by: Incyte Genomics, Inc.	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.nih.gov	
Plate: LNCM1835	row: f	column: 16	High quality sequence stop: 734.	
Location/Qualifiers	1. 957			
/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="IMAGE:5019653"	/clone_id="NIH_MGC_42"	/issue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"	/note="Organ: pancreas; Vector: pOTB; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: This is a NIH_MGC Library."			
BASE COUNT	254 a	255 c	255 g	193 t
ORIGIN				
Query Match	32.6%	Score 664.8;	DB 13;	Length 957;
Best Local Similarity	95.1%	Pred. No. 4.4e-123;		
Matches 719; Conservative	0;	Mismatches 32;	Indels 5;	Gaps 3;
OY	10	GATGAGCGTCTGGGGGGTGCACGAGCAGTTCCTGGGCGAGTGGGCTCAGGGGCTCGA	69	
DB	2	GATGCGCTCGCTGGGGGCTGCACGAGCAGTTCCTGGGCGAGTGGGCTCAGGGGCTCGA	61	
OY	70	ACGAGCCCTACCTCTTGTAGTGTCTCTGCGGGGCGACCGGGCAATCCACGCTGGC	129	
DB	62	ACGGACCTACTCTTGTAGTGTCTCTGCGGGGCGACCGGGCAATCCACGCTGGC	121	
OY	130	GTTGACGTAGGCGCAGCGGCTCGGGGCTGAGATCTCAAGCGCTGACTCCATGCAAGTCTA	189	
DB	122	GTTGACGTAGGCGCAGCGGCTCGGGGCTGAGATCTCTGACGCTGACTCCATGCAAGTCTA	181	
OY	190	TGAAGGCTTACATCATCACAACAAGGTTTCTCCCAAGAGCGAGAATCTGGCGCA	249	
DB	182	TGAAGGCTTACATCATCACAACAAGGTTTCTCCCAAGAGCGAGAATCTGGCGCA	241	
OY	250	CCACATGATCAGCTTTGTGATTCCTCTGTGACCAATTACACAGTGTGGAATTGAGAA	309	
DB	242	CCACATGATCAGCTTTGTGATTCCTCTGTGACCAATTACACAGTGTGGAATTGAGAA	301	
OY	310	TAGAGCAACTGCTGTGATTGA-AGATATATTTGCCCGACAAATTCCTATTTGTTGG	368	
DB	302	TAGAGCAACTGCTGTGATTGAAGATATATTTGCCCGACAAATTCCTATTTGTTGG	361	
OY	369	GAGAACCAATTTACTGTAATCTCGTCCTGGAAGTCTTGTAATCCAAAGCCGC	428	
DB	362	GAGAACCAATTTACTGTAATCTCGTCCTGGAAGTCTTGTAATCCAAAGCCGC	421	
OY	429	AGGAGATGGGACAGAGAAATGATTTGACCCAAAAGTGTGAAAGAGAGATGTGC	488	
DB	422	AGGAGATGGGACAGAGAAATGATTTGACCCAAAAGTGTGAAAGAGAGATGTGC	481	
OY	489	TTGACTTCACAAAGCCTAAGCAGGTGAGCCAGAA-ATGCTGCCAAGCTGCATCA	547	

451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES

source

Location/Qualifiers
 1. 724

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UT-E-EO1-ab-b-20-0-UT"
 /clone_lib="UT-E-EO1"
 /tissue_type="fetal eye"
 /dev_stage="fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pUT73-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UT-E-EO1 is a normalized cDNA library containing the
 following tissue(s): fetal eye. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pUT73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 CCGGTATACC. This library was created for the program, Gene
 Discovery in the Visual System, supported by National Eye
 Institute (NEI)."

BASE COUNT 208 a 132 c 175 g 208 t 1 others
 ORIGIN

Query Match

Best Local Similarity 33.28; Score 677.8; DB 14; Length 724;
 Matches 715; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

OY 1117 CCTGTGACCTGTGTGATCATCATCTGGGATCGCAATGGCGACGACATATA 1176
 |||||||
 Db 1 CCTGTGACCTGTGTGATCATCATCTGGGATCGCAATGGCGACGACATATA 60
 OY 1177 ATCCAAATCCCACTTGAACCACTGAAGAAAGAAAGATTGGACTCAGATGCTCAA 1236
 |||||||
 Db 61 ATCCAAATCCCACTTGAACCACTGAAGAAAGAAAGATTGGACTCAGATGCTCAA 120
 OY 1237 CACCATGAAAGTCAGAGTGTTCCTCCAGACTATTAACAAGACCTTAAGAGGAGATC 1296
 |||||||
 Db 121 CACCATGAAAGTCAGAGTGTTCCTCCAGACATTAACAAGACCTTAAGAGGAGATC 180
 OY 1297 CCCAGGCGAGATGATCAAGAGTCGAAATGAGCGTTTAAGAGACATGCCAGTGGCTT 1356
 |||||||
 Db 181 CCCAGGCGAGATGATCAAGAGTCGAAATGAGCGTTTAAGAGACATGCCAGTGGCTT 240
 OY 1357 TGAAGAGTGGTGGGATCCAGTTCCAGAGGAGGAGGATTTGTCTCCAGTCTGGCC 1416
 |||||||
 Db 241 TGAAGAGTGGTGGGATCCAGTTCCAGAGGAGGAGGATTTGTCTCCAGTCTGGCC 300
 OY 1417 AAAGAGTGGCTATGCGGAATTCCTGATAGCAAGAAAGCTCCACCATTTTCTTTTAT 1476
 |||||||
 Db 301 AAAGAGTGGCTATGCGGAATTCCTGATAGCAAGAAAGCTCCACCATTTTCTTTTAT 360
 OY 1477 GTGGTTTAAAGTCTCAGTCTCTATATATGAAGAAGACAGAGTCTGTGCAGCTCCTTGTG 1536
 |||||||
 Db 361 GTGGTTTAAAGTCTCAGTCTCTATATATGAAGAAGACAGAGTCTGTGCAGCTCCTTGTG 420
 OY 1537 TGCGTGAATGTGTGGAATGATGATGATGAGAAAGA-TTTTCTTTTCTTTGAAACT 1595
 |||||||

Db 421 TGCGTGAATGTGTGGAATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 OY 1596 TAAAGTCTCTATTTTAAAGACACACAGATTCACATTTTATACATGAGATCTCTT 1655
 |||||||
 Db 481 TAAAGTCTCTATTTTAAAGACACACAGATTCACATTTTATACATGAGATCTCTT 540
 OY 1656 TGTGGTGAATCCAGAGATTCAGTCACATCCCTTTAAAGATTTTAAAGTCTGACTCTG 1715
 |||||||
 Db 541 TGTGGTGAATCCAGAGATTCAGTCACATCCCTTTAAAGATTTTAAAGTCTGACTCTG 599
 OY 1716 GCTAAATATATATATTTCCAGATCTTTTGTAGATGATGATGATGATGATGATGAT 1775
 |||||||
 Db 600 GCTAAATATATATATTTCCAGATCTTTTGTAGATGATGATGATGATGATGATGATGAT 659
 OY 1776 ATTTGGATTTTATATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1835
 |||||||
 Db 660 ATTTGGATTTTATATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 718
 OY 1836 GTGAAC 1841
 |||||||
 Db 719 GTGAAC 724

RESULT 10
 B612651 869 bp. mRNA linear EST 18-APR-2001
 B612651
 LOCUS
 DEFINITION
 602640078F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4771338 5',
 mRNA sequence.
 ACCESSION
 B612651
 VERSION
 B612651.1 GI:13664022
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 869)
 AUTHORS
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 CONTACT
 Robert Strausberg, Ph.D.
 COMMENT
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LCM1639 row: k column: 19
 High quality sequence stop: 684.

FEATURES

source

Location/Qualifiers
 1. 869

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4771338"
 /clone_lib="NIH_MGC_61"
 /tissue_type="embryonal carcinoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
 SfiI (ggccgctggcc); Site_2: SfiI (ggccatattggcc);
 Double-stranded cDNA was prepared from cell line RNA. 5'
 and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGCCCTTTGGCC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCGGCGGCGGCGCATG-dt(30)-BN-3'
 (where B = A, C, or G and N = A, C, G, or T). Average
 insert size 1.75 kb (range 0.9-4.0 kb). 15/15 clones
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA). Note: this is a NIH-MGC
 Library."

BASE COUNT 259 a 184 c 206 g 220 t
 ORIGIN

```

|||||
Db 240 GGATCCCTGTGACCAATTCACAGTGTGAGCTTGAAGAACCACTGCTGAT 299
QY 328 TGAAGATATATTTGCCCGAGACAAATTCCTATTGTTGGGAGAACCAATTTCAT 387
Db 300 TGAAGATATATTTGCCCGAGACAAATTCCTATTGTTGGGAGAACCAATTTCAT 359
QY 388 TGAATCTCTGCTGTGAAAGTTCTTGTCAATACCAGCCCCAGAGATGGGCACTGAA 447
Db 360 TGAATCTCTGCTGTGAAAGTTCTTGTCAATACCAGCCCCAGAGATGGGCACTGAA 419
QY 448 AGGATTTGACCCGAAAGTGTGAGTGTGAAAGAGAGATGGTCTTACTTACAAACCCCT 507
Db 420 AGGATTTGACCCGAAAGTGTGAGTGTGAAAGAGAGATGGTCTTACTTACAAACCCCT 479
QY 508 AACCCAGGTGGACCCAGAAATGGGTGCCAGCTGCATGCATGACAAAGCAAGTGGC 567
Db 480 AACCCAGGTGGACCCAGAAATGGGTGCCAGCTGCATGCATGACAAAGCAAGTGGC 539
QY 568 CAGAGCTTGCAAGTTTGTGAAGAAACAGAAATCTCTCATAGTGAATTTTCCATGCTCA 627
Db 540 CAGAGCTTGCAAGTTTGTGAAGAAACAGAAATCTCTCATAGTGAATTTTCCATGCTCA 599
QY 628 ACATAGGAAAGAGTGTGTGCTCCCTTGAGAGTCTCTGAAGTTCTTAACCTTGCAT 687
Db 600 ACATAGGAAAGAGTGTGTGCTCCCTTGAGAGTCTCTGAAGTTCTTAACCTTGCAT 659
QY 688 CCTTGGCTTCATGCTGACAGCAGTTCTAGATGAGCGCTTGATGAAGAGTGGATGA 747
Db 660 CCTTGGCTTCATGCTGACAGCAGTTCTAGATGAGCGCTTGATGAAGAGTGGATGA 719
QY 748 CATGC 752
Db 720 CTGTC 724

```

```

RESULT 8
LOCUS A1133396 725 bp mRNA linear EST 11-NOV-1999
DEFINITION HA2003 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
ACCESSION A1133396
VERSION A1133396.1 GI:6360712
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 725)
AUTHORS Yu, Y., Zhang, C., Luo, L., Ouyang, S., Li, W., Wu, J., Zhou, S., Liu, M.
and He, F.
Expression profile analysis of a human fetal liver cDNA library
JOURNAL Unpublished (1998)
COMMENT Contact: Yongtao Yu
Department of Hematology
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Tel: 0086-10-68159479
Fax: 0086-10-68214653
Email: yyt48@yahoo.com
FEATURES
Location/Qualifiers
1..725
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human fetal liver cDNA library"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="MC1061/P3"
/note="Vector: pCDNA1"

```

BASE COUNT 220 a 177 c 124 g 204 t

Query Match 33.7%; Score 688.2; DB 9; Length 725;

Best Local Similarity 99.3%; Pred. No. 1e-127; Matches 712; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

```

QY 1274 AAGAACTTAAGAGAGAGGAT-CCCGAGGCGAAGATCAAGAGCTGAATACACGCT 1332
Db 717 AAGAACTTAAGAGAGAGGATCCCGAGGCGAAGATCAAGAGCTGAATACACGCT 658
QY 1333 TTAAGACATATGTCACAGTGGCCTTTGGAAGGTGGGGATCCAGTTCAGAGGAGGG 1392
Db 657 TTAAGACATATGTCACAGTGGCCTTTGGAAGGTGGGGATCCAGTTCAGAGGAGGG 598
QY 1393 GTATGTTTGTCTCCCACTGTGGGCAAGAGAGTGTATGCGGAATTTCTGTATGCGAAA 1452
Db 597 GTATGTTTGTCTCCCACTGTGGGCAAGAGAGTGTATGCGGAATTTCTGTATGCGAAA 538
QY 1453 AAGCTCCACCATTTTCTTTGATGATGTTTAAAGTCTCAGTCTCTATAATAGAAAC 1512
Db 537 AAGCTCCACCATTTTCTTTGATGATGTTTAAAGTCTCAGTCTCTATAATAGAAAC 478
QY 1513 AGCAGCTTGTGACCTCTGTGTGCTGATGTGTGTAATGATGTAGTTCAGAGAA 1572
Db 477 AGCAGCTTGTGACCTCTGTGTGCTGATGTGTGTAATGATGTAGTTCAGAGAA 418
QY 1573 GCATTTTCTTTCTTTGAACTTAAGGTTCTATTTAAAGCAGACAGATTCCACA 1632
Db 417 GCATTTTCTTTCTTTGAACTTAAGGTTCTATTTAAAGCAGACAGATTCCACA 358
QY 1633 TTTTATACATGAGAGATCTTGTGTGATATCCAGAGATGAGTGCATCCCTTAAA 1692
Db 357 TTTTATACATGAGAGATCTTGTGTGATATCCAGAGATGAGTGCATCCCTTAAA 299
QY 1693 GAAGTTTATGTCCTGACCTCTGCTAAATTTATCTAATTTCCAGATGCTTTTGTAGATG 1752
Db 298 GAAGTTTATGTCCTGACCTCTGCTAAATTTATCTAATTTCCAGATGCTTTTGTAGATG 239
QY 1753 ACTGAGTATTTGTGAGCCACATATTTGGAGTCTTGAATTTGAGTGAATGGCAGAGAAAG 1812
Db 238 ACTGAGTATTTGTGAGCCACATATTTGGAGTCTTGAATTTGAGTGAATGGCAGAGAAAG 179
QY 1813 GCCATCTCCATGAGATGATTAAGTGAACCAACTAGTTCGGAATTTCTACAGAGAGG 1872
Db 178 GCCATCTCCATGAGATGATTAAGTGAACCAACTAGTTCGGAATTTCTACAGAGAGG 119
QY 1873 AGGGAATCAGACAGAGAGAGTGTGACATAGAGCTTGAAGACCAAGCTTTGAATTTG 1932
Db 118 AGGGAATCAGACAGAGAGAGTGTGACATAGAGCTTGAAGACCAAGCTTTGAATTTG 59
QY 1933 CGAGCTGCTATGTGTGATTTATATCACTGCTGCTTTCTTATTTGAGTTACAAATCT 1989
Db 58 CGAGCTGCTATGTGTGATTTATATCACTGCTGCTTTCTTATTTGAGTTACAAATCT 2

```

```

RESULT 9
LOCUS BM721352 724 bp mRNA linear EST 01-MAR-2002
DEFINITION UI-E-EO1-a1b-20-0-UI-r1 UI-E-EO1 Homo sapiens cDNA clone
ACCESSION BM721352
VERSION BM721352.1 GI:19041207
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 724)
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa

```

University of Iowa


```

/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/Note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
U1-CF-EN1 is a normalized cDNA library containing the
following tissue(s): Primary Lung Cystic Fibrosis
Epithelial Cells. The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT73-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CTCGTACAGT.
TAG_L1B-U1-CF-EN1
TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
6hr to LPS 24h
TAG_SEO=CTGCTCAGGT"
BASE COUNT      231 a      172 c      116 g      224 t
ORIGIN
Query Match      35.6%; Score 727.4; DB 14; Length 743;
Best Local Similarity 99.7%; Pred. No. 1.4e-135;
Matches 739; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1301 GGGCAGATGATCAAGAGCTGAATGACGCTTTAAGAGACATGTCAGTGGCCTTTGGA 1360
DB 743 GGGCAGATGATCAAGAGCTGAATGACGCTTTAAGAGACATGTCAGTGGCCTTTGGA 684
QY 1361 AAGGTGGTGGGATCCAGTTCAGAGAGGAGGATGTTGTCCTCCAGTCTGGGCAAG 1420
DB 683 AAGGTGGTGGGATCCAGTTCAGAGAGGAGGATGTTGTCCTCCAGTCTGGGCAAG 624
QY 1421 GAGTGTATGCGGAATTCCTGATAGCAGAAAAGCTCCACCAATTTCTTTGATGTGG 1480
DB 623 GAGTGTATGCGGAATTCCTGATAGCAGAAAAGCTCCACCAATTTCTTTGATGTGG 564
QY 1481 TTTTAAGTCTCAGCTCTCTATATAGAAAACAGACAGTCTTGTACAGTCTGTGGC 1540
DB 563 TTTTAAGTCTCAGCTCTCTATATAGAAAACAGACAGTCTTGTACAGTCTGTGGC 504
QY 1541 TGATGTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1600
DB 503 TGATGTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 444
QY 1601 GTTCTATTTATTAAGACAGACAGATTCACATTTTATACATGAGATCTTCTTGTGG 1660
DB 443 GTTCTATTTATTAAGACAGACAGATTCACATTTTATACATGAGATCTTCTTGTGG 384
QY 1661 TGAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1720
DB 383 TGAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 325
QY 1721 AATTATCTAATTTCCAGATGCTTTTGTAGATGATGATGATGATGATGATGATGATG 1780
DB 324 AATTATCTAATTTCCAGATGCTTTTGTAGATGATGATGATGATGATGATGATGATG 265
QY 1781 GAGTGTAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1840
DB 264 GAGTGTAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 205
QY 1841 CCAAACTAGATTTCTGGAATTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1900
DB 204 CCAAACTAGATTTCTGGAATTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 145
QY 1901 TAGACTTGAAGACCAAGATTGGAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1960
DB 144 TAGACTTGAAGACCAAGATTGGAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 85
QY 1961 CTCGTCTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2020

```

```

DB 84 CTCGTCTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 25
QY 2021 ATTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2041
DB 24 ATTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4
RESULT 7
LOCUS BE315223
DEFINITION BE315223 725 bp mRNA linear EST 26-OCT-2000
601141778F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3141643 5',
mRNA sequence.
ACCESSION BE315223
VERSION BE315223.1 GI:9145717
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 725)
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: Image.lnl.gov
Plate: L1CML1 row: k column: 20
High quality sequence stop: 720.
FEATURES
source
1..725
Location/Qualifiers
1..725
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3141643"
/clone_1b="NIH_MGC_9"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: ovary; Vector: pORF; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      186 a      175 c      197 g      167 t
ORIGIN
Query Match      33.8%; Score 690.6; DB 10; Length 725;
Best Local Similarity 97.9%; Pred. No. 3.3e-128;
Matches 710; Conservative 0; Mismatches 14; Indels 1; Gaps 1;
QY 28 TGACGAGCAGATTCCTGAGGAGTGGGCTGAGGGGCTGACAGGAGCCCTACCTGTG 87
DB 1 TGACGAGCAGATTCCTGAGGAGTGGGCTGAGGGGCTGACAGGAGCCCTACCTGTG 60
QY 88 AGTATTCCTGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 147
DB 61 AGTATTCCTGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
QY 148 GCTGGGGGAGATGCTGACGCTGATCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAG 207
DB 121 GCTGGGGGAGATGCTGACGCTGATCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
QY 208 CACCAACAGGTTTCTGCCCAAGAGAGAGATGCGCGGACACACATGATGATGATGATG 267
DB 181 CACCAACAGGTTTCTGCCCAAGAGAGAGATGCGCGGACACACATGATGATGATGATG 239
QY 268 GGATCCTCTTGTGACCAATTACAGAGTGTGAGCTTCAAGAAATAGAGCACTGCTGAT 327

```



```

RESULT 5
BM800217      1088 bp      mRNA      linear      EST 05-MAR-2002
LOCUS        AGENCOURT 6416157 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5531248
DEFINITION   5', mRNA sequence.
ACCESSION    BM800217
VERSION      BM800217.1 GI:19117040
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE    1 (bases 1 to 1088)
AUTHORS      NIH-MGC http://mhc.nci.nih.gov/.
TITLES       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cga@bbs.femail.nih.gov
              Tissue Procurement: ATCC
              CDNA Library Preparation: Life Technologies, Inc.
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.llnl.gov
              Plate: L1AM12212 row: j column: 17
              High quality sequence stop: 651.
FEATURES
Source
1..1088
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5531248"
/clone_lib="NIH-MGC_71"
/tissue_type="leiomiosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
BASE COUNT   318 a      218 c      255 g      294 t      3 others
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

1339 GACATGTCAGATGCGCTTTGGAAAGTGTGGGAGATCCAGTTCAGAGGAGGGATATGT 1398
|||||
501 GACATGTCAGATGCGCTTTGGAAAGTGTGGGAGATCCAGTTCAGAGGAGGGATATGT 560
|||||
1399 TTGCTCTCCAGTCTGGGCAAGAGAGTGTATGCCGAAATTCCTCTCATAGCAGAAAGCTC 1458
|||||
561 TTGCTCTCCAGTCTGGGCAAGAGAGTGTATGCCGAAATTCCTCTCATAGCAGAAAGCTC 620
|||||
1459 CCACATTTTCTTTGATGTGGTGTAAAGCTCAGCTTCTCTAATAGAAACAGCAG 1518
|||||
621 CCACATTTTCTTTGATGTGGTGTAAAGCTCAGCTTCTCTAATAGAAACAGCAG 680
|||||
1519 TCTGTGAGCTCTTGTGTGGCTGATGTCTGGAATGATGATGATGAGCAATTT 1578
|||||
681 TCTGTGAGCTCTTGTGTGGCTGATGTCTGGAATGATGATGATGAGCAATTT 740
|||||
1579 TTTTCTTCTTGAACCTTAAGTGTCTATTTAAAGCAGACAGATTCACATTTT 1638
|||||
741 TTTTCTTCTTGAACCTTAAGTGTCTATTTAAAGCAGACAGATTCACATTTT 798
|||||
1639 TACATGAGATCTCTTGTGTGATATACAGATGATGATGATGATGATGATGATGAT 1698
|||||
799 TACATGAGATCTCTTGTGTGATATACAGATGATGATGATGATGATGATGATGAT 858
|||||
1699 TTATGT-CCCTGACTCT-GGCTAAATATCTAATTTCA 1736
|||||
859 TTATGTCCCTGACTCTGGCTAAATATCTAATTTCA 898
|||||
RESULT 6
BM975520/c      743 bp      mRNA      linear      EST 21-MAR-2002
LOCUS        UI-CF-EN1-acw-c-07-0-UI-s1 UI-CF-EN1 Homo sapiens cDNA clone
DEFINITION   UI-CF-EN1-acw-c-07-0-UI 3', mRNA sequence.
ACCESSION    BM975520
VERSION      BM975520.1 GI:19593111
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE    1 (bases 1 to 743)
AUTHORS      Bonaldo, M.F., Lennon, G., and Soares, M.B.
TITLES       Normalization and subtraction: two approaches to facilitate gene
              discovery
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      97044477
COMMENT      Contact: McCray, PB
              McCray Lab
              University of Iowa
              2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
              Tel: 319 356 4866
              Fax: 319 356 7171
              Email: paul-mccray@uiowa.edu
              Tissue Procurement: Dr. M. J. Welsh, University of Iowa
              CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
              CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: Researchers may obtain clones from Research
              Genetics (www.resgen.com).
              The following repetitive elements were found in this cDNA
              sequence: 1-56, >AT-richlow_complexity (matched complement)
              Seq primer: M13 FORWARD
              POLY-A-yes.
FEATURES
Source
1..743
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-acw-c-07-0-UI"
/clone_lib="UI-CF-EN1"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial
Cells"

```

```

Db 567 TCTGACAGGCGAGAAAGGCTCCACCATTTCTTTTGATGTGTTTAAAGTCTCAGCTTC 508
QY 1499 TCTATATATATAAACAGAGGCTCTTGACGCTCTTGCTGCTGCTATGCTGTGAATGA 1558
Db 507 TCTATATATATAAACAGAGGCTCTTGACGCTCTTGCTGCTGCTATGCTGTGAATGA 448
QY 1559 TGTAGTTCAGAAAGCATTTTCTTTCTTTGAACCTTAAAGGCTTATATTAAAGCA 1618
Db 447 TGTAGTTCAGAAAGCATTTTCTTTCTTTGAACCTTAAAGGCTTATATTAAAGCA 388
QY 1619 GCACAGATTCACATTTTATATACATGAGATCTTCTTTGCTGGAATACAGAGATTGACT 1678
Db 387 GCACAGATTCACATTTTATATACATGAGATCTTCTTTGCTGGAATACAGAGATTGACT 328
QY 1679 GCATCCCTTTAAAGAAGTTTATGTCCTGACCTGCTGCTGCTAAATTTATTTCCAGA 1738
Db 327 GCATCCCTTT-AAAAGAGTTTATGTCCTGACCTGCTGCTGCTAAATTTATTTCCAGA 269
QY 1739 TGTCTTTGTAGTACGATGATTTTGTGAGCCACATTTGGAGTCTAGATTGAGTG 1798
Db 268 TGTCTTTGTAGTACGATGATTTTGTGAGCCACATTTGGAGTCTAGATTGAGTG 209
QY 1799 AATGGCAGAAAGGCGCTCTCCATTCATGATGATTAAGTGAACCAACTAGTCTCGAA 1858
Db 208 AATGGCAGAAAGGCGCTCTCCATTCATGATGATTAAGTGAACCAACTAGTCTCGAA 149
QY 1859 TTCTACGAGAAAGAGGATTCAGACCTGAGAGAGCTGTGACATGAGCTGAACCA 1918
Db 148 TTCTACGAGAAAGAGGATTCAGACCTGAGAGAGCTGTGACATGAGCTGAACCA 89
QY 1919 GACTTTAAATTTGCGAGCTGCTGATGAGTATATACATGCTGCTCTTTTATTTGA 1978
Db 88 GACTTTAAATTTGCGAGCTGCTGATGAGTATATATATTACTGCTGCTTTTATTTGA 29
QY 1979 GTTACAAATCTATATTTTATTTGA 2002
Db 28 GTTACAAATCTATATTTTATTTGA 5

RESULT 4
Bi758117 799 bp mRNA linear EST 25-SEP-2001
LOCUS 603023811F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5194347 5'
DEFINITION mRNA sequence.
ACCESSION Bi758117
VERSION Bi758117.1 GI:15749695
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 799)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgaabs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L14M1486 row: e column: 04
            High quality sequence start: 2
            location/Qualifiers
            1. 799
FEATURES
            source
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5194347"
            /clone_lib="NIH_MGC_114"

```

```

/lab_host="DH10B"
/note="Organ: brain; Vector: PCMV-SPORT6; Site:1: NotI;
Site:2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb. Insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Institut fuer Genetik, Research Genetics tracking code 019. Note:
this is a NIH-MGC library."
BASE COUNT 207 a 188 c 220 g 184 t
ORIGIN
Query Match 37.3%; Score 760.8; DB 13; Length 799;
Best Local Similarity 99.4%; Pred. No. 2.9e-142;
Matches 795; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
QY 12 TGGCGTCCGTGGCGGCTGACAGACAGATTCCTGTGGCAGTGGCTCAGGGGCTGCAAC 71
Db 1 TGGCGTCCGTGGCGGCTGACAGACAGATTCCTGTGGCAGTGGCTCAGGGGCTGCAAC 60
QY 72 GGACCTTACCTCTTGTGTGATTTCTCGGGCCACGGGACCCGCAATCCACGCTGGCT 131
Db 61 GGACCTTACCTCTTGTGTGATTTCTCGGGCCACGGGACCCGCAATCCACGCTGGCT 120
QY 132 TGCAGCTAGGCGACGGGCTGGGGGTGAGATTCGTCAGCGTCACTCCATGAGGCTATG 191
Db 121 TGCAGCTAGGCGACGGGCTGGGGGTGAGATTCGTCAGCGTCACTCCATGAGGCTATG 180
QY 192 AAGGCTAGACATCATCACAACAAGTTTCTGCCCAAGACAGAGATTCGCGGACCC 251
Db 181 AAGGCTAGACATCATCACAACAAGTTTCTGCCCAAGACAGAGATTCGCGGACCC 240
QY 252 ACATGATCAGCTTTGTGGATCTCTTGTGACCAATTAACAGTGTGGACTTCAGAAATA 311
Db 241 ACATGATCAGCTTTGTGGATCTCTTGTGACCAATTAACAGTGTGGACTTCAGAAATA 300
QY 312 GAGCAACGCTGCTGATTTGAAGATATATTGGCCGACGAAATTCCTATTGTTGGGAG 371
Db 301 GAGCAACGCTGCTGATTTGAAGATATATTGGCCGACGAAATTCCTATTGTTGGGAG 359
QY 372 GAACCAATTTATACATTGAATCTGCTGTGGAAGTCTTGTCAATACCAAGCCGAG 431
Db 360 GAACCAATTTATACATTGAATCTGCTGTGGAAGTCTTGTCAATACCAAGCCGAG 419
QY 432 AGATGGGCACTGAGAAAGTATTAACGGAAGTGAAGCTTTGAAAGAGAGATGCTTGG 491
Db 420 AGATGGGCACTGAGAAAGTATTAACGGAAGTGAAGCTTTGAAAGAGAGATGCTTGG 479
QY 492 TACTTCAACAAAGGCTTAAGCCAGGTGACCCAGAAATGGCTGCCAAGCTGCATCAGATG 551
Db 480 TACTTCAAC-AAGGCTTAAGCCAGGTGACCCAGAAATGGCTGCCAAGCTGCATCAGATG 538
QY 552 ACAACGCAAAAGTGGCCAGAGAGCTTGGCAAGTTTGAAGAAACAGAGATCTCATAGTG 611
Db 539 ACAACGCAAAAGTGGCCAGAGAGCTTGGCAAGTTTGAAGAAACAGAGATCTCATAGTG 598
QY 612 AATTTCATCATGTCACATACGGAAGAAGTGTGTCCCTTGGAGAGTCTCTGAGT 671
Db 599 AATTTCATCATGTCACATACGGAAGAAGTGTGTCCCTTGGAGAGTCTCTGAGT 658
QY 672 TCTTAACCCCTTGGATCCTTGGCTCATGCTGACAGGAGCTCTAGATGAGGCTTGG 731
Db 659 TCTTAACCCCTTGGATCCTTGGCTCATGCTGACAGGAGCTCTAGATGAGGCTTGG 718
QY 732 ATAAAGAGGTGATGATGCTTGTGCTGCTGCTT-GGAGAGCTAAGAGATTTTCAC 790
Db 719 ATAAAGAGGTGATGATGATGCTTGTGCTGCTGCTGCTGAGAGAACTAAGAGATTTTCAC 778
QY 791 AGAGCTTATATCAGAGAA 810
Db 779 AGAGCTTATATCAGAGAA 798

```

OY	1962	GCATCGCCTTAAGAGGTATCTGATCTCCGAAGTGGGAGGAGTCTGTTCTTCAACCTGTT	1021
Db	964	AACCTAAGAGATATGCCCGGAAACAGAACCGATGGGTTAAACCGCTTTTGTAGC-----	1018
OY	1032	CTTGAATCGTGCAAGTTTCATCCAGGGGCACAAGCCCTACAGCCACTCCAATTAAGATG	1081
Db	1019	-----AGGTCAAAACCTACAGCCATGCGCACTGAAGATG	1052
OY	1082	CCATCAATGAAGCTGAGAACCAAGAAAGTTATCAACCTGTGTGACCTCTGTATGGAATC	1141
Db	1053	GCATCAATGAAGAGGAACAAGAAAGTTAACACATGTGTGACCTCTGTGACCGGATC	1112
OY	1142	ATCATTTGGGAGATCCGGAATTTGGGACACGACATAAATCCAAATCCACTTGAACAACGTG	1201
Db	1113	ATCATTTGGGAGACCGGGAATTTGGGACACACATTTAGAAATCCAAATCTCACTTGCACCAATTG	1172
OY	1202	AAGAAAGAAGAAAGATTGGACTCAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTC	1261
Db	1173	AAGAAAGAAGAAAGTGTGGACTTAAACCTGTCACTGTCCACAGAAATCAAGTAAATTC	1232
OY	1262	CCAGACTTAACAAGAAACCTTAAGGAAAGGATCCCGAGGGCAGAAATGATCAAGAGCTG	1321
Db	1233	CCAGACTGTGACCCGGAAACGCAATCAGAGGGGGAATCCTCGGGGCACACAAATCAAGAGCTG	1292
OY	1332	AAATCCAGCGTTTAAGAGACATGTCAGTGGCCCTTTGGAAGGTGGTGGGATTCAGATTC	1381
Db	1293	AAGCCAGATGTTTGAAGACATGCTCATGTGGCTTTTGCGA-GAGTGGGATTCAGAGCC	1351
OY	1382	AGGAGGAGAGGGTATGTTTGTCTCCA-GTCTGGGCAAAAGAGTGCATGCGGAATCTTC	1440
Db	1352	AGGAGGAGAGGGAGTGTGCTCTCCACGCGCTGAGAGTGAGGAATCTGCGGC-----	1402
OY	1441	TGCATAGCAAGAAAGCTCCACACTTCTTTTGAATGATGGTTTAAAGTCAACGTTCTC	1500
Db	1403	-----AGAGAGCCCAACCACTTCTTTCATTTCTG-----TCTGTGTC	1441
OY	1501	TATAATAGAAACAGCAGGTCCTTGTCAGCTCCTTGTGTGGCTGATGTGTGAAATGATG	1560
Db	1442	TGCAGTGGAAACAGCAGGCTT-TCAGTCTCTTGTGTGCTGTGTGTGGTAATGATG	1500
OY	1561	TAGTTCAGAAAGCATTTTCTTTTCTTTGAACCTTAAAGGTTCTATATTTAAAGCAGC	1620
Db	1501	TAGTTCAGAGTGGGA--TTTTTTTCTTTGAACCTTAAAGGTTTATTTTAAAGCAGAGC	1557
OY	1621	ACAGATTCACATTTTATATACATGAGGATCTTCTTGTGTGTAATAACAGGATTCGCTG	1680
Db	1558	ACAGATTCGACATTTTCTACTGAGAGATCTTTTATGTGTGTAAATCCAGATTCAGTGC	1617
OY	1681	ATCCCTTTAAAGAAAGTTTATGTCCCTGACCTGCTGCTAAATTTATCAATTTCCAGATG	1740
Db	1618	AT-CCTTTAAAGACGTTCTGTCCCTGGCGCTGCTGCTAAATATAGCTGTTTCCAGATG	1676
OY	1741	CTTTGTGATGATCAGTGAAGTATTTGTGAGCCCAATTTGGAGTCTAGATTGGAGTAA	1800
Db	1677	CTTTGTGTGATGATCAGTGAAGTATTTGAAGCCGAGTCAGAGAGCTCTGAGATCGAGAGCGC	1736
OY	1801	TGGCAGGAAGAGGCCATCCCATTCAGATGATTAAGTGAACCAACTAGTCTCGGAATT	1860
Db	1737	AGAAAGGAGTAG-----TGCAGAAAGATGATTAAGCAAACTCTCCACACTCTGTGAAT-	1789
OY	1861	CTACAGAGAAGAGGAAATCAGATCAGAGAAAGCTGTGACATAGGAGCTTGAAGAC-CAAG	1919
Db	1790	---CTACAGAGAAGAGGGGTCAGGCTGAGGTGTGTGAACTGTGATCTTGAACATCAAAAG	1846
OY	1920	ACTTTGAATTTTGCAGAGTCTCATGTG	1947
Db	1847	ACACAGAGTCTGCTGAGCTGCTCTGTG	1874

DEFINITION	AL530574 LTI.NFL001.NBC4 Homo sapiens CDNA clone CS0DD008YN11 3 prime-mRNA sequence.
ACCESSION	AL530574
VERSION	AL530574.1 GI:12794067
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 987)
TITLE	L1.W.B., Gruber,C., Jesse,J. and Polayes,D.
JOURNAL	Full-length cDNA libraries and normalization
COMMENT	Unpublished (2001)
FEATURES	Contact: Genoscope
source	Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seqlife@genoscope.cns.fr , Web : www.genoscope.cns.fr . Location/Qualifiers
	1..987
	/organism="Homo sapiens"
	/db_xref="taxon:3606"
	/clone="CS0DD008YN11"
	/clone_1lb="LTI.NFL001.NBC4"
	/sex="male"
	/tissue_type="neuroblastoma cells"
	/lab_host="DH10B"
	/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand CDNA was primed with a NotI-D1190(dt) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9600 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com "
BASE COUNT	281 a 230 c 180 g 284 t 12 others
ORIGIN	
Query Match	46.1%; Score 941.4; DB 9; Length 987;
Best Local Similarity	97.2%; Pred. No. 2e-178;
Matches 956; Conservative 10; Mismatches 17; Indels 1; Gaps 1,	
Oy	1019 GCTCTTGAATCGTGCCAAGTTTCATCCAGGGCCACACGCCCTACAGCCACTCCATAAAG 1078
Db	987 GCTCTTGAATCGTGCCAAGTTTCATCCAGGGCCACACGCCCTACAGCCACTCCATAAAG 928
Oy	1079 ATGCATCAATGAAGCTGAGAACAGAAAGAACTTATACCTGTGTGACTCTGTGTGCA 1138
Db	927 ATGCATCAATGAAGCTGAGAACAGAAAGAACTTATACCTGTGTGACTCTGTGTGCA 868
Oy	1139 ATCATCATTTGGGGATCGCGAATGGGACAGCGACATATAATCCCAATCCCATTTGAACCA 1198
Db	867 ATCATCATTTGGGGATCGCGRATGGGGCGGCGCACATATAATCCCAATCCCATTTGAACCA 808
Oy	1199 CTGAAGAAAGAGAGAGATTGGACTCAGTCTGTCTCAACACATAGAAAGTCAAGTGT 1258
Db	807 CTGAAGAAAGAGAGAGATTGGACTCAGTCTGTCTCAACACATAGAAAGTCAAGTGT 748
Oy	1259 TCCCCAGACTATACAAAGAACTTAAGGGAAGGGATCCCCAGGGCAGATATATCAAG 1318
Db	747 TCCCCAGACCAATACAAAGAACTTAAGGGAAGGGATCCCCAGGGCAGATATATCAAG 688
Oy	1319 CTGAATCAGAGGTTTAAGAGACATGTCAGAGGGCTTTGAAAGGTGGTGGGATCCAG 1378
Db	687 CTGAATCAGAGGCTTTAATAAACTTTGTCAGAGGGCTTTGAAAGGTGGTGGGATCCAG 628
Oy	1379 TTCAGAGGAGGAGGTATATTGTCTCCAGTCTGGCCAAAGAGAGTGCATGCGGAATTC 1438
Db	627 TTCAGAGGAGGAGGTATATTGTCTCCAGTCTGGCCAAAGAGAGTGCATGCGGAATTC 568
Oy	1439 TCTGCATAGCAGAAAGCTCCCACTTTCTTTGATGTGTGTTTAAAGTTCACGTTTC 1498

QY 548 CATGCAAAACGCAAGTGGCCAGAGCTTCAGATTTTGAAGAAACAGAACTCTCTCAT 607
 |||||
 Db 552 CATGCAAAACGCAAGTGGCCAGAGCTTCAGATTTTGAAGAAACAGAACTCTCTCAC 611
 QY 608 AGTGAATTTCTCCATCCATCAACATGAGAGAGGTGTGTGCTCCCTTGGAGTCCCTG 667
 |||||
 Db 612 AGTGAATTTCTCCATCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 671
 QY 668 AAGTTCTCTAACCTTTCATCTCTTGGCTTCATGCTGACCAAGAGCTTCATGAGAGAG 727
 |||||
 Db 672 AAGTTCTCTAACCTTTCATCTCTTGGCTTCATGCTGACCAAGAGCTTCATGAGAGAGAG 731
 QY 728 TTGATTAAGAGAGGTGAGTACATCTGCTGCTGCTGCTTGGAGAGACTAAGAGATTTT 787
 |||||
 Db 732 TTGATTAAGAGAGGTGAGTACATCTGCTGCTGCTGCTTGGAGAGAGCTTGAAGAGGT 791
 QY 788 CACAGACCTATATATCAGAGAAATGTTTCGAAATATAGCCAGAGCTATCAATCATCTATC 847
 |||||
 Db 792 CACAGACCTATATATCAGAGAAATGTTTCGAAATATAGCCAGAGCTATCAATCATCTATC 851
 QY 848 TTCCATCAATTTGCTTCAAGAGATTTCAAGATCTGATCACTGAGAGAGAAATGACACA 907
 |||||
 Db 852 TTCCATCAATTTGCTTCAAGAGATTTCAAGATCTGATCACTGAGAGAGAAATGACACA 911
 QY 908 CTGGAGACTAGTAACCAAGCTTTTAAAGAAAG----- 938
 |||||
 Db 912 CTGGAGACTAGTAACCAAGCTTTTAAAGAAAGGTATGAGGCTGTAAACAACTAATTAAG 971
 QY 938 -----GACCTGCTCC 949
 Db 972 AGATATGCCCCGGAACAGACCCGATGGTTTAAACCCGCTTTTGGACAGAGCTGGGCC 1031
 QY 950 AATGTCCTCCCTGTATGCTTGAAGTATCTGATGCTTCGAGTGGAGAGAGCTGTT 1009
 |||||
 Db 1032 AGTGTCCCTCCGCTTATGAGTGTAGAGATCTGATGTTTCAAGTGGAGAGAGCTGTT 1091
 QY 1010 CTGGAACCTGCTTGAAGTGTGCAAGTTCATCCAGAGGCGCACAGAGCTTACAGCCT 1069
 |||||
 Db 1092 CTGGAACCTGCTTGAAGTGTGCAAGTTCATCCAGAGGCGCACAGAGCTTACAGCCT 1151
 QY 1070 CCAATTAAGATGCTATACATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1129
 |||||
 Db 1152 CCAATTAAGATGCTATACATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1211
 QY 1130 TGTGATGCAATCATATTGGGAGTGGAGATGGGAGGCGACATTAATCCAAATCCAC 1189
 |||||
 Db 1212 TGTGATGCAATCATATTGGGAGTGGAGATGGGAGGCGACATTAATCCAAATCCAC 1271
 QY 1190 TTGAACCAATGAGAGAGAGAGAGAGATGGAGTGCATGCTGTCAACCATTAAGAGT 1249
 |||||
 Db 1272 TTGACCAATGAGAGAGAGAGAGAGATGGAGTGCATGCTGTCAACCATTAAGAGT 1331
 QY 1250 CAGAGTGTTCCTCCAGACTATTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1309
 |||||
 Db 1332 CAGAGTGTTCCTCCAGACTATTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1391
 QY 1310 GATCAAGAGCTGAAGTGAAGGCTTTAAGAGAGATGCTGAGTGGCTTTGAAGAGTGTG 1369
 |||||
 Db 1392 AATCAAGAGCTGAAGTGAAGGCTTTAAGAGAGATGCTGAGTGGCTTTGAAGAGTGTG 1450
 QY 1370 GGGATCCAGTTCAGAGAGAGAGAGAGTGTGTCTCCCA-GTCTGGGCAAGAGAGTGTG 1428
 |||||
 Db 1451 GGGATCCAGTTCAGAGAGAGAGAGAGTGTGTCTCCCA-GTCTGGGCAAGAGAGTGTG 1510
 QY 1429 TGGGAAATTTCTGATAGAGAGAGAGAGTGTGTCTCCCA-GTCTGGGCAAGAGAGTGTG 1488
 |||||
 Db 1511 GGC-----AGAGGCCCCACCATCTTCTTTATTTCTGTGCT----- 1546
 QY 1489 TCTCAGTTCCTATATAGAGAGAGAGAGTGTGTCTGAGCTCTTGTGTGCTGATGTGT 1548
 |||||
 Db 1547 -----GTGGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1599
 QY 1549 CTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1608

Db 1600 CTGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1656
 |||||
 QY 1609 ATTAAAGCAGACAGATTCACATTTTATACATGAGATCTTCTTGGAGATAC 1668
 |||||
 Db 1657 TTGAAGCAGAGCAGATTCACATTTTCTACTGAGATCTTTTGAAGATGATAC 1716
 QY 1669 AGATTTGACTGATCCCTTTTAAAGAGATTTTATGCTCCCTGACTGCTGCTTAAATATCT 1728
 |||||
 Db 1717 AAGATTCAGTGCAT-CTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1775
 QY 1729 AATTTCCATATGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1788
 |||||
 Db 1776 CGTTCCATATCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1835
 QY 1789 GATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1848
 |||||
 Db 1836 CTGAG 1889
 QY 1849 GTTCTCGAATTTCTACAG 1908
 |||||
 Db 1890 GCTCTGTGAT-----CTACAG 1945
 QY 1909 GAAGAC-CAAGACTTTGAATTTGCGAGCTGCTCATGTG 1947
 |||||
 Db 1946 GAAGACTTAAAG 1985
 |||||
 RESULT 2
 AK010176
 LOCUS
 DEFINITION
 Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310075G14:homolog to TRNA ISOPTERYLYPROPHOSPHATE TRANSFERASE, full insert sequence.
 AK010176
 VERSION
 AK010176.1 GI:12845432
 KEYWORDS
 HTG; CAP trapper.
 SOURCE
 Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone:1b:RIKEN full-length enriched mouse cDNA library
 clone:2310075G14.
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.
 REFERENCE
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE
 99279253
 PUBMED
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE
 2049374
 PUBMED
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsui, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE
 20530913
 PUBMED
 11076861
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:27:18 ; Search time 3842.56 Seconds
(without alignments)
8602.339 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041
Sequence: 1 CTGCCATAGATGCGTCGCG.....TTTACAGAGAAAAA 2041

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthm:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pin:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1128.4	55.3	2061	11 AK003556	AK003556 Mus muscu
2	1069.6	52.4	1951	11 AK010176	AK010176 Mus muscu
3	941.4	46.1	987	9 AL530574	AL530574 AL530574
4	760.8	37.3	799	13 B1758117	B1758117 603023811
5	733.6	35.9	1088	14 BM600217	BM600217 AGENCOURT
6	727.4	35.6	743	14 BM975520	BM975520 UI-CF-EMI

Result No.	Score	Query Match	Length	ID	Description
7	690.6	33.8	725	10 BE315223	BE315223 601141778
8	688.2	33.7	725	9 A1133396	A1133396 HA2003 Hu
9	677.8	33.2	724	14 BM721352	BM721352 UI-E-E01
10	666.2	32.6	869	12 BG612651	BG612651 602640078
11	664.8	32.6	957	13 B1161353	B1161353 602865759
12	662.8	30.8	772	12 BG034660	BG034660 602300022
13	606.8	29.7	624	14 B0003256	B0003256 UI-H-E11
14	599.8	29.4	639	10 AV759288	AV759288 AV759288
15	594.8	29.1	711	12 BG506209	BG506209 601860153
16	594	28.9	783	13 B1222788	B1222788 602941574
17	589	28.9	601	9 A1207688	A1207688 HA3066 Hu
18	578.2	28.3	922	9 AL530575	AL530575 AL530575
19	574.4	28.1	578	14 BM847727	BM847727 K-EST0127
20	570	27.9	643	10 AM953634	AM953634 EST365704
21	569.6	27.9	758	13 B1820461	B1820461 603034775
22	546.8	26.8	747	12 BG702127	BG702127 602684032
23	546.4	26.8	559	9 A1718450	A1718450 a556807.x
24	544.2	26.7	716	13 B1101508	B1101508 602887307
25	541.8	26.5	806	12 BE965936	BE965936 601659823
26	539.4	26.4	724	12 BF207826	BF207826 601862530
27	535	26.2	554	9 A1749267	A1749267 a41e07.x
28	530.4	26.0	543	9 A1193562	A1193562 qe70f07.x
29	526.2	25.8	836	12 BG575745	BG575745 602597234
30	524.4	25.7	562	12 BG495857	BG495857 602540248
31	515.4	25.3	528	12 BF064275	BF064275 7157h03.x
32	511.2	25.0	963	13 B1462398	B1462398 603203761
33	509.2	24.9	946	12 BG288423	BG288423 602388087
34	504.8	24.7	645	10 BB656365	BB656365 BB656365
35	500.4	24.5	513	10 AW104415	AW104415 xD86502.x
36	495.8	24.3	510	10 AW263011	AW263011 x153f07.x
37	490.2	24.0	518	9 A1459111	A1459111 t165b11.x
38	487.4	23.5	500	9 A1192804	A1192804 qe67912.x
39	479	23.5	490	9 A1127851	A1127851 qe42a03.x
40	476	23.3	487	9 A1337829	A1337829 qe29f07.x
41	470.4	23.0	491	12 BF433525	BF433525 7934604.x
42	466.8	22.9	546	10 AV607062	AV607062 AV607062
43	465.4	22.8	593	10 AM444470	AM444470 UI-H-B13
44	459	22.5	862	12 BF142859	BF142859 601787926
45	458.6	22.5	676	10 BE614088	BE614088 601503823

ALIGNMENTS

RESULT 1
AK003556
LOCUS
DEFINITION
Mus musculus 18 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:111007017, homolog to TRNA
ISOPENTENYLPIROPHOSPHATE TRANSFERASE, full insert sequence.
ACCESSION
AK003556
VERSION
AK003556.1 GI:12834290
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (strain:C57BL/6J) 18 days embryo cDNA to mRNA,
clone:111007017.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
JOURNAL
MEDLINE
PUBMED
11042159

NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Amster, Rotstein & Ebenstein
 STREET: 90 Park Avenue
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10016
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch 1.44 Mb storage disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: Word Processor (ASCII)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/075,904
 FILING DATE: May 11, 1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/247,901
 FILING DATE: May 23, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Bogosian, Elizabeth A
 REGISTRATION NUMBER: 39,911
 REFERENCE/DOCKET NUMBER: 96700/475
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 697-5995
 TELEFAX: (212) 286-0854 or 286-0082
 TELEX: TMX 710-581-6766
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 50341
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE:
 DESCRIPTION: L5 shuttle phasmid sequence
 HYPOTHETICAL: NO
 ANTI-SENSE:
 FRAGMENT TYPE:
 ORIGINAL SOURCE:
 ORGANISM: L5 mycobacteriophage
 STRAIN:
 INDIVIDUAL ISOLATE:
 DEVELOPMENTAL STAGE:
 HAPLOTYPE:
 TISSUE TYPE:
 CELL TYPE:
 CELL LINE:
 ORGANELLER:
 IMMEDIATE SOURCE:
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT:
 FEATURE:
 NAME/KEY:
 LOCATION:
 IDENTIFICATION METHOD:
 OTHER INFORMATION:
 PUBLICATION INFORMATION: No. 5994137e
 AUTHORS:
 TITLE:
 JOURNAL:
 VOLUME:
 PAGES:
 DATE:
 DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO:
 US-09-075-904-1

Query Match	1.8%;	Score 36.4;	DB 2;	Length 50341;
Best Local Similarity	49.0%;	Pred. No. 10;		
Matches 97; Conservative	0;	Mismatches 101;	Indels 0;	Gaps 0;

QY	40	TCGTTGGGAGTGGGCTCAGGGGGCTCGCAAGGAGACCCACCTCTTTGATGAGATTCGG	99
Db	2478	TCCTCTGTGATGTTCTCTGTGCGCGGGCTTGAGCGGCCCACTACTCTCTTGGCGACCTGC	24119
QY	100	GGCCACGGGCGCCAGGCAATCCACGCTGGCGTTGCAGCTAGGCCAGCGGCTGGCGGTGA	159
Db	2418	GGCGAAGTTCAGCCTCGACCTCTGCTGGCCCGCTGGCGCTTCGGCGAGCTTCCTGTGGAGA	2359
QY	160	GATCGTCAGCGCTGACTCCATGCGAGTCTATATGAGGCTTAGACATCATCACCACACAGGT	219
Db	2358	CCCTCGAGGCTCCCTGAGCCCTCGTAGTTGAGGATCTCGAGACCCCTCGCACCGCCCCACT	22399
QY	220	TTTCGCCAAGACAGCAGAG	237
Db	2298	TCGTGGCCGACAGCCGGG	2281

Search completed: April 21, 2003, 16:30:04
Job time : 3773.69 secs

STATE: CA
COUNTRY: US
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/053,374A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: COOK, ROBERT R.
REGISTRATION NUMBER: 31,602
REFERENCE/DOCKET NUMBER: A-514
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1689 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 132..971
US-09-053-374A-4

Query Match 1.8%; Score 36.4; DB 4; Length 1689;
Best Local Similarity 66.7%; Pred. No. 1.7;
Matches 52; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1964 CTGCTTTCTTATGATGACAACTATTTTATTGAACTTTAAATGAAGAAAAATT 2023
DB 1544 CTACATTTCTATGTGTGAGAGCTTTATTATTATTAAATTAATTAATTTTA 1603
QY 2024 TACAAGAAAAAATAA 2041
DB 1604 AAATGAAAAAATAA 1621

RESULT 14
US-08-247-901C-1/C
Sequence 1, Application US/08247901C
Patent No. 5750384
GENERAL INFORMATION:
APPLICANT: Jacobs et al
TITLE OF INVENTION: L5 SHUTTLE PHASMIDS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rothstein & Edenstein
STREET: 90 Park Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/247,901C
FILING DATE: May 23, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/057,531
FILING DATE: April 29, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Elizabeth A
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 96700/273
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 50341
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: L5 shuttle phasmid sequence
HYPOTHETICAL: No
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: L5 mycobacteriophage
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
ORGANELLE:
IMMEDIATE SOURCE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION: No. 5750384e
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-247-901C-1

Query Match 1.8%; Score 36.4; DB 1; Length 50341;
Best Local Similarity 49.0%; Pred. No. 10;
Matches 97; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 40 TCCTGTGGGAGTGGGCTCAGAGGGGCTGCAAGGACCTTCTTGTAGTATTCGG 99
DB 2478 TCCTGTGGTGTAGTCTCTGCGCGGCTGAGCGGCCAGTCACCTTGTGCGC 2419
QY 100 GGGCAGGGGCAACGGCAATGCAAGCTGCGGCTTGCAGTACGCGCGCGGCGA 159
DB 2418 GGGCAAGTTCAGCTCGACCTGCTGCGCCCTGCGCTTCGCGAGCTTCTGCGG 2359
QY 160 GATGTCAGCGCTGCTCATGTCAGAGGCTTATGAGGCGCTAGACATCATCACAAGGT 219
DB 2358 CCTGAGCGTCTCGAGACCTCGACCTGCTGCGCCCTGCGCTTCGCGAGCTTCTGCGG 2359
QY 220 TTCTGCCCAAGACAGAG 237
DB 2298 TCGTGGCCGACAGCGCGG 2281

RESULT 15
US-09-075-904-1/C
Sequence 1, Application US/09075904
Patent No. 5994137
GENERAL INFORMATION:
APPLICANT: Jacobs, et al
TITLE OF INVENTION: L5 SHUTTLE PHASMIDS

```

; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

```

```

Query Match      1.8%; Score 37.4; DB 4; Length 2346;
Best Local Similarity 64.4%; Pred. No. 1;
Matches 56; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

```

```

QY 1955 TATCTAGCTGCTCTTATGAGTTACAAATCTATATTTTATGAGTTAAATAA 2014
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2197 TTTTATGATGATATTTGTTATTTCCAGCAATAATTTGTAAGTTGCAAAAAA 2256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2015 GAAAAAATTACAGAAAAA 2041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2257 AAAAAAAAAAAAAAAAAA 2283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 12
US-07-847-010-12
; Sequence 12, Application US/07847010
; Patent No. 5693495
; GENERAL INFORMATION:
; APPLICANT: Breiteneder, Helmo
; APPLICANT: Reikstorfer, Arnold
; APPLICANT: Valenta, Rudolf
; APPLICANT: Hofmann, Sommergruber, Karin
; APPLICANT: Breitenbach, Michael
; APPLICANT: Kraft, Dietrich
; APPLICANT: Rumpold, Helmut
; APPLICANT: Scheiner, Otto

```

```

; APPLICANT: Ebner, Christof
; APPLICANT: Ferrel, Fatima
; TITLE OF INVENTION: Allergens of Alder Pollen and
; TITLE OF INVENTION: Applications Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/847,010
; FILING DATE: 01-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jones III, Harry C
; REGISTRATION NUMBER: 20,280
; REFERENCE/DOCKET NUMBER: 6530-010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 742 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: hazel (Corylus sp.)
; IMMEDIATE SOURCE:
; LIBRARY: POLLEN FROM ALLERGEN AB, ENGELHOLM, SWEDEN
; US-07-847-010-12

```

```

Query Match      1.8%; Score 36.6; DB 1; Length 742;
Best Local Similarity 56.1%; Pred. No. 0.94;
Matches 69; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

```

```

QY 1919 GACCTTGAATTTGCGAGCTGCTGAGTTATATTCACAGCTGCTTCTATTTGA 1978
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 604 GATGTTTAACTCGGAAGATGCAAACTGAATGTTGTAATTTTGCATATATACAA 663
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1979 GTTACAAATCTAATTTTATTTGAGTTAAATTAAGAAAAAATTACAGAAAAA 2038
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 664 TAATGGAAGATTAACATCATTTGAGTTCAAAAAAAAAAAAAAAAAAAAAA 723
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2039 AAA 2041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 724 AAA 726
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 13
US-09-053-374A-4
; Sequence 4, Application US/09053374A
; Patent No. 6462177
; GENERAL INFORMATION:
; APPLICANT: YEN, KWANG-MU
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: ONE AMGEN CENTER DRIVE
; CITY: THOUSAND OAKS

```

[illegible]


```
Patent No. 6465229
GENERAL INFORMATION:
APPLICANT: Rafalski, Antoni J.
APPLICANT: Fader, Gary M.
APPLICANT: Cahoon, Rebecca E.
TITLE OF INVENTION: Plant Caffeoyl-CoA O-Methyltransferase
FILE REFERENCE: B01284 US NA
CURRENT APPLICATION NUMBER: US/09/452,239
CURRENT FILING DATE: 1999-12-01
EARLIER APPLICATION NUMBER: 60/110,594
EARLIER FILING DATE: 1998-December-02
NUMBER OF SEQ ID NOS: 50
SOFTWARE: Microsoft Office 97
SEQ ID NO 21
LENGTH: 980
TYPE: DNA
ORGANISM: Glycine max
US-09-452-239-21

Query Match
Best Local Similarity 1.9%; Score 37.8; DB 4; Length 980;
Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1929 TTTCGAGTCTCATGCTGAGTATTAATCACTGCTCTTCTATTGATTACAATC 1988
DB 867 TTTCAGAGCTTGTGTGTTCTGTGACGCTAGAAATTTAGCTGAAGATCAATC 926

QY 1989 TAATATTTTATGAGTTAAATTAAGAAAAATTTACAAGAAAAAAA 2041
DB 927 AACATTATCTTATTTAAAAAAAATTTAAAAAAAATTTAAAAAAA 979

RESULT 9
US-09-453-323-1
Sequence 1, Application US/09453323
Patent No. 6441272
GENERAL INFORMATION:
APPLICANT: Ye, Zheng-Hua
TITLE OF INVENTION: MODIFICATION OF LIGNIN CONTENT AND COMPOSITION IN
PLANTS
FILE REFERENCE: 235,00090101
CURRENT APPLICATION NUMBER: US/09/453,323
CURRENT FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: 60/110,676
PRIOR FILING DATE: 1998-12-02
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1507
TYPE: DNA
ORGANISM: Nicotiana tabacum
US-09-453-323-1

Query Match
Best Local Similarity 1.9%; Score 37.8; DB 4; Length 1507;
Matches 60; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1945 GTGTGAGTATTAATCACTGCTCTTCTATTGAGTTACAATCTATTTTATGAG 2004
DB 1366 GTTGTGAGCTGATTCGCTGTTATTCCTCAATTCCTCCCTACGAGATATGACAGATG 1425

QY 2005 TTTAAATTAAGAAAAATTTACAAGAAAAAAA 2041
DB 1426 ATAAAAAAAATTTAAAAAAAATTTAAAAAAA 1462

RESULT 10
US-09-149-476-316
Sequence 316, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
```

```
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,560
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
```

EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match Best Local Similarity 1.9% Score 38 DB 4 Length 1605;
Matches 53; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1964 CTGCTCTTATGAGTTACAAATCTATTATTTTGAAGTTAATAAGAAAAATT 2023
DB 1506 CTGCTTAATAATAGACTGTGTAATTAATGTTGAGTCTGCAAAAAAAA 1565
QY 2024 TACAAGAAAAAAA 2041
DB 1566 AAAAAAAA 1583

RESULT 8
US-09-452-239-21
Sequence 21, Application US/09452239

DB 4122 GGTGGAAACGAGATCATCGACGCTGCGAGAGATATACCGTGACATGACATCG 4181
QY 207 TCACCAACAAGCTTCTGCGCCAGACGAG 235
DB 4182 GCACAGCAAGACTTGCGCGACTATCAG 4210

RESULT 6

US-08-974-691-7
Sequence 7, Application US/08974691
Patent No. 6225103

GENERAL INFORMATION:

APPLICANT: Keolisch, Gerald
APPLICANT: Lin, Xinh
TITLE OF INVENTION: Cloning and Characterization of Napsin
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1201 W. Peachtree
St.
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,691

FILING DATE: 20-NOV-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/031,196

FILING DATE: 20-NOV-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/046,126

FILING DATE: 09-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Pabst, Patrea L.

REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: OMFR 166

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404-873-8794

TELEFAX: 404-873-8795

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1910 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

US-08-974-691-7

Query Match

Best Local Similarity 61.1%; Score 40.8; DB 4; Length 1910;
Matches 66; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1934 GAGCTGCTATGTGTGATATATACAGTCTCTTTCTATTGAGTTACAATCTATAT 1993
DB 1796 GTGAACCGCATCGCAGCGGTGTCGCGAGCTCCGTACCCAGTAAATCCACTAT 1855
QY 1994 TTTTATGAAGTTAAATAGAAAAATTACAGAAAAA 2041
DB 1856 TTCCATTGAAAAA 1903

RESULT 7

US-09-149-476-187
Sequence 187, Application US/09149476
Patent No. 6420526

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 186 Human Secreted proteins

FILE REFERENCE: P2002pl

CURRENT APPLICATION NUMBER: US/09/149,476

FILING DATE: 1998-09-08

EARLIER APPLICATION NUMBER: PCT/US98/04493

EARLIER FILING DATE: 1998-03-06

EARLIER APPLICATION NUMBER: 60/040,162

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,333

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/038,621

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,626

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,334

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,336

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,163

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/047,600

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,615

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,597

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,502

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,633

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,563

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,617

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,618

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,503

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,592

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,581

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,584

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,500

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,587

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,492

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,598

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,613

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,582

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,596

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,612

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,632

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,601

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/043,580

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,568

QY 442 TGAGAAAGTATGACCGAATAAGTATGATGCTTCTGACTTCA 498
DB 381 AATGAAACATTAATTAATTAAGTATGACCTTAATTAATTAAGTCTCA 440
QY 499 CAAAGCCCTAAGCAGGCTGAGCCAGAAATGCTGCAAGCTGCATCCATGACAAAGC 558
DB 441 CGAATATTATGCTCATCGCAAGAAATGACCAAGATATCATCTTAATACAGAAA 500
QY 559 CAAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGAA 599
DB 501 AAGAGTGTGCGACCAATAGATATTAATTAAGAAACAAAA 541

RESULT 2
US-09-103-840A-1/c
Sequence 1, Application US/09103840A

Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 2.9%; Score 58.4; DB 4; Length 4411529;
Best Local Similarity 53.7%; Pred. No. 5.1e-05;
Matches 175; Conservative 0; Mismatches 136; Indels 15; Gaps 2;

QY 87 TAGGATCTCTGCGGCGCCAGCCGCAATCCAGCTGCGCTTGCAGTACGCGCAGC 146
DB 3040756 TGCGGATTAATCGGACCGACCGCGCCGCAAGTCCAGTGGCTTGCAGCTGCCCGCGC 3040697
QY 147 GGCTCGGCGGT-----GAGATCGTCAAGCGCTGACTCCATGCAAGGTCTATGAG 194
DB 3040696 GGCTCGGCGCGCAGGCTGAGCGTGAATCGTGAACGCCGACGATCGAGTGTATGCGC 3040637
QY 195 GCGTAGCATATCATCAACAAGGTTTGTGCCAAGACAGAGATTCGCGGCGCACACA 254
DB 3040636 GCATGACATCGGACGCGCAAGTGCCTCAAGCGCGCGTGGCATCCCGCATCATC 3040577
QY 255 TGACAGCTTGTGATGCTCTGTGACCAATTAACAGTGGTGAATTCAGAAATAGAG 314
DB 3040576 AACTCGATGCTCTGATGATCAACCAAC---GGACCGTGGCGCTACCAAGCGGCGC 3040520
QY 315 CAAGTCTCTGATTAAGATATATTTGCCGAGACAAATTCCTATTGTGTGGAGAA 374
DB 3040519 CCGGCGGAGCATGAGCGCATCGCAGCCGCGGAGCGGTGCGTGTGGCGGCT 3040460
QY 375 CCAATTATTACATTGAATCTGCTC 400
DB 3040459 CGATGCTGTATGTCCAATCCCTGCTC 3040434

RESULT 3
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEFFLINGER, F.
APPLICANT: FALKNER, F. G.

TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 2.6%; Score 52.2; DB 1; Length 7218;
Best Local Similarity 3.1%; Pred. No. 0.0001;
Matches 12; Conservative 222; Mismatches 155; Indels 0; Gaps 0;

QY 1073 ATAAAGATGCCATTAATGAAGCTGAGAAAGAGAGAGTATACCTGTGACCTCTGT 1132
DB 1449 AGAAGATTTGTACRR 1390
QY 1133 GATCAATCATCATTTGGGATCGCAATGGCGACGCGCATTAATAATCCAAATCCACTG 1192
DB 1389 RRR 1330
QY 1193 AACCAATGAAGAAAGAAAGAAAGATTTGACTGATGCTCAACACCATAGAAAGTCAG 1252
DB 1329 RRR 1270
QY 1253 AGTGTTCCTCCAGACTATTAACAAGACCTAAAGGAGGATCCCGAGCGCAATGAT 1312
DB 1269 RRR 1210
QY 1313 CAAGAGCTGAATGCAAGCTTTAAGAGACATGCTCCAGTCTGGGCAAGAGTGGTGGG 1372
DB 1209 RRR 1150
QY 1373 ATCCAGTTGAGGAGGAGGATGTTGTCTCCAGTCTGGGCAAGAGTGGTATGCG 1432
DB 1149 RRR 1090
QY 1433 GAATTCCTGATAGCAAGAAAGTCCCA 1461
DB 1089 RRRRRRRRRRRRRRRRRRRRRRRRRATCGCA 1061

1	65	3.2	582	4	US-09-134-001C-1334	Sequence 1334, App
2	58.4	2.9	4411	582	US-09-133-840A-1	Sequence 1, Appl
3	52.2	2.6	7218	1	US-08-203-463-14	Sequence 14, Appl
4	47.8	2.3	846	4	US-08-936-165A-65	Sequence 65, Appl
5	46.6	2.3	5884	4	US-09-921-017B-187	Sequence 497, App
6	40.8	2.0	1910	4	US-08-927-661-7	Sequence 7, Appl
7	38	1.9	1605	4	US-09-149-476-187	Sequence 187, App
8	37.8	1.9	980	4	US-09-452-239-21	Sequence 21, Appl
9	37.4	1.9	1507	4	US-09-455-323-1	Sequence 1, Appl
10	37.4	1.8	2025	4	US-09-149-476-316	Sequence 316, App
11	37.4	1.8	2346	4	US-09-149-476-193	Sequence 193, App
12	36.6	1.8	742	1	US-07-847-010-12	Sequence 12, Appl
13	36.4	1.8	1689	4	US-09-053-374A-4	Sequence 4, Appl
14	36.4	1.8	50341	1	US-08-247-901C-1	Sequence 1, Appl
15	36.4	1.8	50341	2	US-09-075-901C-1	Sequence 1, Appl
16	36.4	1.8	53297	4	US-09-426-436-1	Sequence 1, Appl
17	36.4	1.8	53297	4	US-08-705-557-1	Sequence 1, Appl
18	36.4	1.8	72604	4	US-09-266-992-7	Sequence 7, Appl
19	36.4	1.8	72604	4	US-09-657-474-7	Sequence 7, Appl
20	36.2	1.8	236	4	US-09-453-702B-165	Sequence 165, App
21	36.2	1.8	2372	4	US-09-145-391-1	Sequence 1, Appl
22	36.2	1.8	3602	4	US-09-402-929-1	Sequence 1, Appl
23	36	1.8	2007	3	US-08-747-221B-36	Sequence 36, Appl
24	36	1.8	2007	3	US-08-747-221B-38	Sequence 38, Appl
25	36	1.8	2007	4	US-09-005-051-36	Sequence 36, Appl
26	36	1.8	2007	4	US-09-005-051-38	Sequence 38, Appl
27	36	1.8	2700	3	US-09-315-861-1	Sequence 1, Appl

28	3.6	1.8	2700	4	US-09-398-386A-43	Sequence 43, Appl
29	3.6	1.8	5852	1	US-07-867-105-2	Sequence 2, Appl
30	35.8	1.8	6702	3	US-08-987-439-2	Sequence 2, Appl
31	35.6	1.7	817	4	US-08-885-469-1	Sequence 1, Appl
32	35.6	1.7	817	4	US-09-625-918-1	Sequence 9, Appl
33	35.6	1.7	1930	4	US-09-724-864-9	Sequence 9, Appl
34	35.6	1.7	3812	4	US-09-784-316-1	Sequence 1, Appl
35	35.6	1.7	5852	1	US-07-867-106-2	Sequence 1, Appl
36	35.6	1.7	6611	1	US-08-880-443-1	Sequence 2, Appl
37	35.4	1.7	6611	1	US-08-457-459-1	Sequence 1, Appl
38	35.4	1.7	6611	1	US-08-555-678-1	Sequence 1, Appl
39	35.4	1.7	6611	5	PCR-US55-02255-1	Sequence 1, Appl
40	35.4	1.7	4581	1	US-08-524-757-1	Sequence 1, Appl
41	35.2	1.7	1737	1	US-08-202-056-4	Sequence 4, Appl
42	35.2	1.7	1737	1	US-08-076-093A-3	Sequence 3, Appl
43	35.2	1.7	1737	1	US-08-701-266-3	Sequence 3, Appl
44	35.2	1.7	1737	1	US-08-284-556-3	Sequence 3, Appl
45	35.2	1.7	1737	2	US-08-805-478-3	Sequence 3, Appl

ALIGNMENTS

```

RESULT 1
US-09-134-001C-1334
: Sequence 1334, Application US/091344001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GNC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: CURRENT FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 1334
: LENGTH: 582
: TYPE: DNA
: ORGANISM: Staphylococcus epidermidis
: US-09-134-001C-1334

```

Query Match	Score	DB	Length
3.2%	65	4	582

Best Local Similarity	48.0%;	Pred. No. 5.2e-09;							
Matches	250;	Conservative	0;	Mismatches	265;	Indels	6;	Gaps	2

QY	82	TCCTGTGATGTAATTTCTCGGGGCGACAGGGGCACCCGCAAAATCCACGCTGGCGTTGCAGTACG	14
Db	24	TTTATATGTTATTTGTAGTGTCCAACTGGTTCACAGTAAACACGATGAATTAAGTATTAAGATTTGC	83
QY	142	CCAGCGGCTCGGCGGTAGATGCTGCACGGCTACCTCATGATGAGGTCTATGAAGCCCTAGA	201
Db	84	TAAAAATTTTAATGGAAGAAATTTATTACGGGAGATTCAATGCAAGGTCTATCAAGGAAATGGA	143
QY	202	CATCATCACAACAAGGTTTCTGCCCCACAGACAGATCTGCCGGCACACATGATCAG	261
Db	144	TATGTGTACAGCAAAATTTCAACTGAAGAAATGGAAGTATACCATTAATATGATAGA	203
QY	262	CTTTGTGATCCCTCTTGTGACCAATTACACAGTGTGTGACTTCAGAAATTAAGCACTGC	321
Db	204	TATTTTGGCTCCAGATGCTTCCTTTCTTGATAT - - GAATTTAAAAAAGGCGACGAAAA	260
QY	322	TCGTGATGAGATATATTTCGCCGAGCAAAATTCCTATTGTGTGGAGAGAACCATTA	381
Db	261	ATATATTAAAGATATTACTGGAAGAGCGAAGTGCCTATTATAGCAGAGGAACAGAGACT	320
QY	382	TTACATTTGAATCTCTGCTGTGAAAAGTTCTGTCAATACCACACCCGAGAGATGGGCAC	441
Db	321	ATATTATCAATCTCTCTTATACAACTATGCTTTTGAAGATGAAATCCATATCTGAAGTTAA	380

```
RESULT 15
US-09-835-232-6/c
; Sequence 6, Application US/09835232
; Patent No. US20020098489A1
; GENERAL INFORMATION:
; APPLICANT: Leder, Philip
; APPLICANT: Leder, Benjamin
; TITLE OF INVENTION: FORMIN-2 NUCLEIC ACIDS AND POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 00383/052002
; CURRENT APPLICATION NUMBER: US/09/835,232
; PRIORITY FILING DATE: 2001-04-12
; PRIORITY FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 180216
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1) (180216)
; OTHER INFORMATION: n = A,T,C or G
US-09-835-232-6

Query Match      2.0%; Score 39.8; DB 10; Length 180216;
Best Local Similarity 53.5%; Pred. No. 22;
Matches 83; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1885 TGAGGACCTGTGACATGAGCTTGAAGACCAAGACTTTGAATTTGGAGCTGCTCAT 1944
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102428 TTAGAAACCACTGACACATATATGACTTTTCAAGATATTTAAGTTATATAGCTCCT 102369
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1945 GTGTGAGTTATATACAGCTGCTCTTCTATGAGTTACAATCTATATTTTATTGAAG 2004
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102368 GCGTGACTTATTTAATCTTGATTTTGTCTGAGGTAAGCTAATATTTTATCATACNA 102309
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2005 TTTAATAAAGAAAAAATTTACAGAAAAA 2039
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102308 GATAAATTTTAAAGAAAAATTTGGAAGTATGAA 102274
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: April 21, 2003, 18:47:42
Job time : 370.611 secs

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1394)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-410

Query Match 2.1%; Score 42.4; DB 10; Length 1401;
Best Local Similarity 79.0%; Pred. No. 0.11;
Matches 49; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 1980 TTCAATCTATATTTTATGAGTTAATAAGAAAAATTTACAGAAAAAAA 2039
DB 1295 TTTAAATAATATTTTGAACCTTMAAAAAA 1354

QY 2040 AA 2041
DB 1355 AA 1356

RESULT 12
US-09-878-574-11368
Sequence 11368, Application US/09878574
Patent No. US20020110548A1
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 11368
LENGTH: 257
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: 701064366H1
US-09-878-574-11368

Query Match 2.0%; Score 41.4; DB 10; Length 257;
Best Local Similarity 52.6%; Pred. No. 0.065;
Matches 90; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 645 GTGTCCTTGGAGCTCTGAGTTCTTAACCTTGACCTTGCTTCATGCTG 704
DB 12 GTGTCCTTGGAGCTCTGAGTTCTTAACCTTGACCTTGCTTCATGCTG 71
QY 705 ACCAGCAGTTCTAGATGACCGCTTGATAGAGGTGATGACATCTTGCTGGC 764
DB 72 TTAAGGAGCTCTGAGTGGGGGTGATATTCCTACAGTCAACAAGCTTGCTGGT 131
QY 765 TCTTGGAGCACTAGAGATTTTCACAGCGCTATATCAAGAAATGTTT 815
DB 132 TTGATAAGAGAGAGAGAGCTTGATCAGAGTTTCATCGAAATATGTTT 182

RESULT 13
US-09-878-574-11381
Sequence 11381, Application US/09878574
Patent No. US20020110548A1
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21

PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 11381
LENGTH: 263
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: 701064379H1
US-09-878-574-11381

Query Match 2.0%; Score 41.4; DB 10; Length 263;
Best Local Similarity 52.6%; Pred. No. 0.065;
Matches 90; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 645 GTGTCCTTGGAGCTCTGAGTTCTTAACCTTGACCTTGCTTCATGCTG 704
DB 10 GTGTCCTTGGAGCTCTGAGTTCTTAACCTTGACCTTGCTTCATGCTG 69

QY 705 ACCAGCAGTTCTAGATGACCGCTTGATAGAGGTGATGACATCTTGCTGGC 764
DB 70 TTAAGGAGCTCTGAGTGGGGGTGATATTCCTACAGTCAACAAGCTTGCTGGT 129

QY 765 TCTTGGAGCACTAGAGATTTTCACAGCGCTATATCAAGAAATGTTT 815
DB 130 TTGATAAGAGAGAGAGAGCTTGATCAGAGTTTCATCGAAATATGTTT 180

RESULT 14
US-09-938-842A-1242
Sequence 1242, Application US/09938642A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Mang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 1242
LENGTH: 2526
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-1242

Query Match 2.0%; Score 39.8; DB 9; Length 2526;
Best Local Similarity 49.8%; Pred. No. 1;
Matches 101; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 408 TTCTTGCAATPACCAACCCAGAGATGGGCACTGAGAAAGTATGACCGAAATGG 467
DB 1211 TACTAGCAGATACAAACACCGCTCAAGATCAAGATTCAGAACTGAAAGATCTGG 1270
QY 468 AGCTGAAAGAGAGATGCTTCTTACTTCACAAAGCCTTAAGCAGGTGACCCAGAAA 527
DB 1271 ATTCTGAAAGAGAAACCGCAATCGAAAGTTAAACGAAAGACACAGAAAGCTTAAGACT 1330
QY 528 TGGCTGCCAAGCTGCATCCATGACATGACAAAGCAAGTGGCCAGAGCTTCAAGTTTGG 587
DB 1331 TGATTAAGAAAGTTGAATCCCATGAAACGTAATTCAGAGAGCAACAAAGCAGTTCTTG 1390
QY 588 AAGAAACGAATCTCTCATAGT 610
DB 1391 AAGCATCTGAGTGTGATAGTACT 1413

RESULT 9
US-09-939-980-65/C
Sequence 65, Application US/09939980
Patent No. US20020082234A1
GENERAL INFORMATION:
APPLICANT: Black, Michael
Burnham, Martin
Hodgson, John
Knowles, David
Lonetto, Michael
Nicholas, Richard
Pratt, Julie
Reichard, Richard
Rosenberg, Martin
Ward, Judith
TITLE OF INVENTION: No. US20020082234A1e1 Prokaryotic Polynucleotides,
Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/939,980
FILING DATE: 27-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/936,165
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 846 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-09-939-980-65
Query Match 2.3%; Score 47.8; DB 10; Length 846;
Best Local Similarity 51.7%; Pred. No. 0.0019;
Matches 109; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
DB 82 TCTTGAGTATCTCGGGGCCACGGGCAACGCAATCAGCGCTGCGTGCAGCTAG 141
DB 214 TATTTAGTAATTTGGGGCAACTGCTTCAGGTAAACAGAGCTTAGCATGAACTCGC 155
DB 142 CCAGCGCTCGGGGAGATGCTGAGCGCTGATCGATGAGGTTATGAAGGCTTAG 201
DB 154 GAAGCTATCATGTGTAATCATGACGCTGATTTATGCAAGTTTACAAACATATGA 95
DB 202 CATCATCAACCAAGGTTTTCGCCAGAGAGAGAAATTCGCCGACACCATGATCAG 261
DB 94 TATTGAACTGCACAGTAACACCTGAAGAAATGATGTATTCACATCATTAATTTGA 35

DB 262 CTTTGTGATCCCTGTGACCAATTACACA 292
DB 34 TATCTTGAATTCCTGATGATACATTTTCAGCA 4
RESULT 10
US-09-925-302-261
Sequence 261, Application US/09925302
Patent No. US20020044941A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA104
CURRENT APPLICATION NUMBER: US/09/925,302
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05918
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 896
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 261
LENGTH: 2952
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (20)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (33)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (35)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-261
Query Match 2.1%; Score 43.4; DB 10; Length 2952;
Best Local Similarity 54.8%; Pred. No. 0.096;
Matches 86; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
DB 1885 TGAGAGAGCTGACATAGACACTTGAAGACCAAGACTTGAATTTGCGAGCTGCTCAT 1944
DB 2795 TGTTCATGTGGTATTTTGTATTTAAAGATCAAAATTTTGCATATGCTATGCTTCT 2854
DB 1945 GGTGTGATTTATATCATCTGCTCTTCTTATGAGTTACAAATCTATTTTATTTGAAG 2004
DB 2855 ATACCTGTGTGAACGCTAGGATCTCATTTGCCCATGTTAATTTTCTAATAATTAAT 2914
DB 2005 TTTAATAAAGAAAAAATTTCACAGAAAAA 2041
DB 2915 TGACAAAAACAAAAA 2951
RESULT 11
US-09-925-301-410
Sequence 410, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 410
LENGTH: 1401
TYPE: DNA


```

RESULT 5
US-09-864-761-26692/c
Sequence 26692, Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-x-1
CURRENT FILING DATE: 2001-05-23
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 26692
LENGTH: 138
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL033527.25
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HELLA, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: RT HIT: AF074918.1, EVALUATE 3.00e-72
OTHER INFORMATION: EST_HUMAN HIT: BE242161.1, EVALUATE 3.00e-72
OTHER INFORMATION: SWISSPROT HIT: Q9KAC3, EVALUATE 2.00e-03
US-09-864-761-26692
Query Match 6.8%; Score 138; DB 10; Length 138;
Best Local Similarity 100.0%; Pred. No. 1e-30; Indels 0; Gaps 0;
Matches 138; Conservative 0; Mismatches 0;

```

QY	185	GCCTGTGAAGGCTTGACATCATCAACCAAGGTTTCTGCCAAGAGAGAGATCTGC	244	
DB	138	GTCTTGAAGGCTTGACATCATCAACCAAGGTTTCTGCCAAGAGAGAGATCTGC	79	
QY	245	CGGCACCAATGATGAGCTTTGTGATCTCTTGTGACCAATTAACAGTGGTGACTTC	304	
DB	78	CGGCACCAATGATGAGCTTTGTGATCTCTTGTGACCAATTAACAGTGGTGACTTC	19	
QY	305	AGAAATAGAGCACTGCT 322		
DB	18	AGAAATAGAGCACTGCT 1		
RESULT 6				
US-09-738-626-2133				
/ Sequence 2133, Application US/09738626				
/ Publication No. US20020197605A1				
/ GENERAL INFORMATION:				
APPLICANT: NAKAGAWA, SATOSHI				
APPLICANT: MIZOGUCHI, HIROSHI				
APPLICANT: ANDO, SEIKO				
APPLICANT: HAYASHI, MIKIRO				
APPLICANT: OCHIAI, KEIKO				
APPLICANT: YOKOI, HARUHIKO				
APPLICANT: TATEISHI, NAOKO				
APPLICANT: SENO, AKIHIRO				
APPLICANT: IKEDA, MASATO				
APPLICANT: OKAZAKI, AKIO				
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES				
FILE REFERENCE: 249-125				
CURRENT APPLICATION NUMBER: US/09/738, 626				
CURRENT FILING DATE: 2000-12-18				
PRIOR APPLICATION NUMBER: JP 99/377484				
PRIOR FILING DATE: 1999-12-16				
PRIOR APPLICATION NUMBER: JP 00/159162				
PRIOR FILING DATE: 2000-04-07				
PRIOR APPLICATION NUMBER: JP 00/280988				
PRIOR FILING DATE: 2000-08-03				
NUMBER OF SEQ ID NOS: 7059				
SOFTWARE: Patentin ver. 3.0				
SEQ ID NO 2133				
LENGTH: 903				
TYPE: DNA				
ORGANISM: Corynebacterium glutamicum				
US-09-738-626-2133				
Query Match	3.7%	Score 75;	DB 9;	Length 903;
Best Local Similarity	54.2%;	Pred. No. 1.8e-11;		
Matches 175; Conservative	0;	Mismatches 145;	Indels 3;	Gaps 1;
QY	78	TACCTCTTGTATGATGATCTCGGGGGCAGCGGACCGGCAATTCACAGCTGGCGTTGCAC	137	
DB	5	TACACCGATCCAGTGGTGGACCCACATCTGGAAAAATCAGTTTGGAAATTGCTC	64	
QY	138	TAGCCAGGAGGCTCGGCGGTGAGATCGTACGCGTACTCCATGACGAGCTATGAAGGCC	197	
DB	65	TAGCCACAGCTTGACGCGTGAAGTAGTAAATGTGATTCATGCAAGCTGTACAAAGCA	124	
QY	198	TAGACATCATCAACCAAGGTTTCTGCCCAAGAGCAGAGAAATCTGCCGACCAACATGA	257	
DB	125	TGACATCGGACGCAAGAGCTGACTGTGAGAAAGCAGCAAGGCAATTCGCGCATCAAC	184	
QY	256	TGAGCTTTGTGATCTCTTGTGACCAATTAACAGTGGTGGACTTCAGAAATAGAGCA	317	
DB	185	TGATGTGTGGAGCTTACCGAAC--TGCCTAGTGGCGAGGTTTCATTCGAGCGC	241	
QY	318	CTGCTCTGATTAAGATATATTGGCCGAGACAAATTCCTATTGTGTGGGAGAGACA	377	
DB	242	TTCGCGAGTGAAGATATATATGTGTCGCGGCAAAACCCCATCTTGTGTGGCGCTCA	301	
QY	378	ATTATTACATTGATCTGCTC 400		

```

; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4310
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-4310

```

Query Match 15.6%; Score 318; DB 9; Length 318;
Best Local Similarity 100.0%; Pred. No. 9e-84;

Matches 318: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1690 AAGAAGTTTATGTCCTGAGTCTGGCTAAATATCTAATTCAGATGCTTTGTAG 1749
    |||
DB 1 AAGAAGTTTATGTCCTGAGTCTGGCTAAATATCTAATTCAGATGCTTTGTAG 60
QY 1750 ATGACGAGTATTTGTGAGCCACATATTGGAGTTCTGATTTGAGTAATGGCAGAA 1809
    |||
DB 61 ATGACGAGTATTTGTGAGCCACATATTGGAGTTCTGATTTGAGTAATGGCAGAA 120
QY 1810 AGGGCATCTCCATGAGATGATTAAAGTACCAACTAGTCTGGAAATTCACAGAGA 1869
    |||
DB 121 AGGGCATCTCCATGAGATGATTAAAGTACCAACTAGTCTGGAAATTCACAGAGA 180
QY 1870 AGGAGGAATCAGACTGAGGAAGCTGTGACATAGACTTGAAGACCAAGACTTTGAAT 1929
    |||
DB 181 AGGAGGAATCAGACTGAGGAAGCTGTGACATAGACTTGAAGACCAAGACTTTGAAT 240
QY 1930 TTGGCAGCTGCATCTGTGAGATTATTCACACTGCTGTTCTTCTATTTGAGTTACAAT 1989
    |||
DB 241 TTGGCAGCTGCATCTGTGAGATTATTCACACTGCTGTTCTTCTATTTGAGTTACAAT 300
QY 1990 ATATTTTATGAGTTT 2007
    |||
DB 301 ATATTTTATGAGTTT 318

```

RESULT 4
US-09-864-761-10050/c

; Sequence 10050, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; GENE EXPRESSION ANALYSIS BY MICROARRAY

```

; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10050
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL033527.25
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
US-09-864-761-10050

```

Query Match 7.1%; Score 145; DB 10; Length 457;
Best Local Similarity 100.0%; Pred. No. 2e-32;

Matches 145: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 183 AGGTCTATGAGGCTAGACATCATCACCACAGAGTTTCTGGCCAGAGAGAGATCT 242
    |||
DB 450 AGGTCTATGAGGCTAGACATCATCACCACAGAGTTTCTGGCCAGAGAGAGATCT 391
QY 243 GCCGGCACCACATGATCAGCTTTGTGATCTCTTGTGACCAATTAACAGAGTGGTACT 302
    |||
DB 390 GCCGGCACCACATGATCAGCTTTGTGATCTCTTGTGACCAATTAACAGAGTGGTACT 331
QY 303 TCAGAAATGAGCACTGCTCTGAT 327
    |||
DB 330 TCAGAAATGAGCACTGCTCTGAT 306

```



```
OY 61 GGGCTGCAAGGAGCCCTACCTCTGTAGTATCTCGGGGGCCAGGGGACCGGCAATTC 120
    |||||||
DB 61 GGGCTGCAAGGAGCCCTACCTCTGTAGTATCTCGGGGGCCAGGGGACCGGCAATTC 120
OY 121 CACGCTGGCGTTGAGCTAGAGCCAGCGGCTCGGGGTGAGATCGTCACGCTGACTCCAT 180
    |||||||
DB 121 CACGCTGGCGTTGAGCTAGAGCCAGCGGCTCGGGGTGAGATCGTCACGCTGACTCCAT 180
OY 181 GCAGCTCTATGAGGCGCTAGACATCATATCCACAAGGTTTCTGCCCAAGACAGAGAT 240
    |||||||
DB 181 GCAGCTCTATGAGGCGCTAGACATCATATCCACAAGGTTTCTGCCCAAGACAGAGAT 240
OY 241 CTGCGGACCCACATGATGATCTTTGTGATCCTCTTGACCAATTCACAGTGTGA 300
    |||||||
DB 241 CTGCGGACCCACATGATGATCTTTGTGATCCTCTTGACCAATTCACAGTGTGA 300
OY 301 CTTCAGAAATAGAGCACTGCTGTGATTAAGATATATTTGCCGAGACAAATTCCTAT 360
    |||||||
DB 301 CTTCAGAAATAGAGCACTGCTGTGATTAAGATATATTTGCCGAGACAAATTCCTAT 360
OY 361 TGTGTGGAGAGCAACCAATTTATACATGAAATCTCTGTGAAAAGTTCTGTCAATAC 420
    |||||||
DB 361 TGTGTGGAGAGCAACCAATTTATACATGAAATCTCTGTGAAAAGTTCTGTCAATAC 420
OY 421 CAAGCCCGAGAGATGGGCACTGAGAAAGTATTGACCGAAAGTGGAGCTGAAGAAG 480
    |||||||
DB 421 CAAGCCCGAGAGATGGGCACTGAGAAAGTATTGACCGAAAGTGGAGCTGAAGAAG 480
OY 481 GGATGCTCTTACTTTCACAACGCGCTTAAGCAGGTGAGCCAGAAATGGCTGCCAAGT 540
    |||||||
DB 481 GGATGCTCTTACTTTCACAACGCGCTTAAGCAGGTGAGCCAGAAATGGCTGCCAAGT 540
OY 541 GCATCCACATACAAAGCCAAAGTGGCCAGAGCTTGCAAGTTTGAAGAAACAGGAT 600
    |||||||
DB 541 GCATCCACATACAAAGCCAAAGTGGCCAGAGCTTGCAAGTTTGAAGAAACAGGAT 600
OY 601 CTCATATGATGATTTTCATCGATCGTCAACATA-CGGAAGAGGTGTGTGCTCC-TTGA 658
    |||||||
DB 601 CTCATATGATGATTTTCATCGATCGTCAACATA-CGGAAGAGGTGTGTGCTCC-TTGA 658
OY 659 GGTCTCTGAGATCTCTTAACCCCTGATCCTTTGG 694
    |||||||
DB 659 GGTCTCTGAGATCTCTTAACCCCTGATCCTTTGG 694
OY 661 GGTCTCTGAGATCTCTTAACCCCTGATCCTTTGG 696
    |||||||
DB 661 GGTCTCTGAGATCTCTTAACCCCTGATCCTTTGG 696

RESULT 2
US-09-764-853-364
; Sequence 364, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P206
; CURRENT APPLICATION NUMBER: US/09/764,853
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 364
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (619)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (625)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (655)
; OTHER INFORMATION: n equals a,t,g, or c
```

```
NAME/KEY: SITE
LOCATION: (673)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (713)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-853-364

Query Match
32.4% Score 661.8; DB 10; Length 717;
Best Local Similarity 98.4%; Pred. No. 2.4e-185;
Matches 685; Conservative 2; Mismatches 7; Indels 2; Gaps 2;

OY 1 CTGCCATAGATGAGGCTCGGCGGCTGCAAGACAGTCTCTGAGGAGGAGGCTGAG 60
    |||||||
DB 1 CTGCCATAGATGAGGCTCGGCGGCTGCAAGACAGTCTCTGAGGAGGAGGCTGAG 60
OY 61 GGGCTGCAAGGAGCCCTACCTCTGTAGTATCTCGGGGGCCAGGGGACCGGCAATTC 120
    |||||||
DB 61 GGGCTGCAAGGAGCCCTACCTCTGTAGTATCTCGGGGGCCAGGGGACCGGCAATTC 120
OY 121 CACGCTGGCGTTGAGCTAGAGCCAGCGGCTCGGGGTGAGATCGTCACGCTGACTCCAT 180
    |||||||
DB 121 CACGCTGGCGTTGAGCTAGAGCCAGCGGCTCGGGGTGAGATCGTCACGCTGACTCCAT 180
OY 181 GCAGCTCTATGAGGCGCTAGACATCATATCCACAAGGTTTCTGCCCAAGACAGAGAT 240
    |||||||
DB 181 GCAGCTCTATGAGGCGCTAGACATCATATCCACAAGGTTTCTGCCCAAGACAGAGAT 240
OY 241 CTGCGGACCCACATGATGATCTTTGTGATCCTCTTGACCAATTCACAGTGTGA 300
    |||||||
DB 241 CTGCGGACCCACATGATGATCTTTGTGATCCTCTTGACCAATTCACAGTGTGA 300
OY 301 CTTCAGAAATAGAGCACTGCTGTGATTAAGATATATTTGCCGAGACAAATTCCTAT 360
    |||||||
DB 301 CTTCAGAAATAGAGCACTGCTGTGATTAAGATATATTTGCCGAGACAAATTCCTAT 360
OY 361 TGTGTGGAGAGCAACCAATTTATACATGAAATCTCTGTGAAAAGTTCTGTCAATAC 420
    |||||||
DB 361 TGTGTGGAGAGCAACCAATTTATACATGAAATCTCTGTGAAAAGTTCTGTCAATAC 420
OY 421 CAAGCCCGAGAGATGGGCACTGAGAAAGTATTGACCGAAAGTGGAGCTGAAGAAG 480
    |||||||
DB 421 CAAGCCCGAGAGATGGGCACTGAGAAAGTATTGACCGAAAGTGGAGCTGAAGAAG 480
OY 481 GGATGCTCTTACTTTCACAACGCGCTTAAGCAGGTGAGCCAGAAATGGCTGCCAAGT 540
    |||||||
DB 481 GGATGCTCTTACTTTCACAACGCGCTTAAGCAGGTGAGCCAGAAATGGCTGCCAAGT 540
OY 541 GCATCCACATGACAACCCAAAGTGGCCAGAGCTTGCAAGTTTGAAGAAACAGGAT 600
    |||||||
DB 541 GCATCCACATGACAACCCAAAGTGGCCAGAGCTTGCAAGTTTGAAGAAACAGGAT 600
OY 601 CTCATATGATGATTTTCATCGATCGTCAACATA-CGGAAGAGGTGTGTGCTCC-TTGA 658
    |||||||
DB 601 CTCATATGATGATTTTCATCGATCGTCAACATA-CGGAAGAGGTGTGTGCTCC-TTGA 658
OY 659 GGTCTCTGAGATCTCTTAACCCCTGATCCTTTGG 694
    |||||||
DB 659 GGTCTCTGAGATCTCTTAACCCCTGATCCTTTGG 694
OY 661 GGTCTCTGAGATCTCTTAACCCCTGATCCTTTGG 696
    |||||||
DB 661 GGTCTCTGAGATCTCTTAACCCCTGATCCTTTGG 696

RESULT 3
US-09-796-692-4310
; Sequence 4310, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THE
; FILE REFERENCE: 2077 001200
; CURRENT APPLICATION NUMBER: US/09/796,692
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 15:55:39 ; Search time 235.611 Seconds
(without alignments)
8702.120 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041

Sequence: 1 CTGCCATAGATGGCGTCCG.....TTTACAGAGAAAAA 2041

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, NA.*

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	663	32.5	717	10	US-09-764-853-166 Sequence 166, App
2	661.8	32.4	717	10	US-09-764-853-364 Sequence 364, App
3	318	15.6	318	9	US-09-796-692-4310 Sequence 4310, App
4	145	7.1	457	10	US-09-864-761-10050 Sequence 10050, A
5	138	6.8	138	10	US-09-864-761-26692 Sequence 26692, A
6	75	3.7	903	9	US-09-738-626-2133 Sequence 2133, App
7	50.4	2.5	8519	10	US-09-070-927A-106 Sequence 106, App
8	49.8	2.4	989	7	US-08-781-986A-800 Sequence 800, App
9	47.8	2.3	846	10	US-09-939-980-65 Sequence 65, App
10	43.4	2.1	2952	10	US-09-925-302-261 Sequence 261, App
11	42.4	2.1	1401	10	US-09-925-301-410 Sequence 410, App
12	41.4	2.0	257	10	US-09-878-574-11368 Sequence 11368, A
13	41.4	2.0	263	10	US-09-878-574-11381 Sequence 11381, A
14	39.8	2.0	2526	9	US-09-938-842A-1242 Sequence 1242, App
15	39.8	2.0	180216	10	US-09-835-232-6 Sequence 6, App
16	39.4	1.9	368	10	US-09-960-352-5552 Sequence 5552, App
17	39.2	1.9	632	9	US-09-992-331-4 Sequence 4, App
18	39.2	1.9	3092	10	US-09-925-302-310 Sequence 310, App
19	38.8	1.9	377	10	US-09-960-352-7419 Sequence 7419, App

ALIGNMENTS

C 20	38.8	1.9	424	10	US-09-960-352-11218	Sequence 11218, A
C 21	38.8	1.9	722	12	US-10-001-879-1074	Sequence 104, App
C 22	38.4	1.9	265	10	US-09-969-373-1074	Sequence 1074, App
C 23	38.4	1.9	454	10	US-09-960-352-8765	Sequence 8765, App
C 24	38.4	1.9	486	9	US-09-832-292-36	Sequence 36, App
25	38.2	1.9	1170	10	US-09-729-674-119	Sequence 119, App
26	38	1.9	1605	9	US-09-809-391-187	Sequence 187, App
27	38	1.9	2821	9	US-09-736-457-1669	Sequence 1669, App
28	38	1.9	2821	9	US-09-902-941-1669	Sequence 1669, App
29	38	1.9	2821	9	US-09-849-626-1669	Sequence 1669, App
30	38	1.9	2821	9	US-10-017-754-1669	Sequence 1669, App
31	37.8	1.9	933	9	US-10-260-877-31	Sequence 31, App
32	37.8	1.9	980	10	US-09-452-239-21	Sequence 21, App
33	37.8	1.9	1231	10	US-09-891-126-3	Sequence 3, App
C 34	37.8	1.9	4006	10	US-09-925-300-580	Sequence 580, App
C 35	37.8	1.9	16472	9	US-10-079-854-199	Sequence 199, App
C 36	37.8	1.9	16472	10	US-09-764-878-199	Sequence 199, App
C 37	37.8	1.9	32249	9	US-10-079-854-202	Sequence 202, App
C 38	37.8	1.9	32249	10	US-09-764-878-202	Sequence 202, App
C 39	37.4	1.8	213	10	US-09-919-580-736	Sequence 736, App
C 40	37.4	1.8	349	10	US-09-770-791-712	Sequence 712, App
C 41	37.4	1.8	373	10	US-09-960-352-836	Sequence 836, App
C 42	37.4	1.8	2025	9	US-09-809-391-316	Sequence 316, App
C 43	37.4	1.8	2346	9	US-09-809-391-193	Sequence 193, App
C 44	37.4	1.8	2972	9	US-10-057-832-1	Sequence 1, App
C 45	37.2	1.8	430	10	US-09-960-352-2920	Sequence 2920, App

```
RESULT 1
US-09-764-853-166
; Sequence 166, Application US/09764853
; Patent No. US20020096072A1
;
GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206
; CURRENT APPLICATION NUMBER: US/09/764,853
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 166
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (619)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (625)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (655)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (673)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (713)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-853-166
Query Match 32.5%; Score 663; DB 10; Length 717;
Best Local Similarity 98.4%; Pred. No. 1.1e-185;
Matches 665; Conservative 3; Mismatches 6; Indels 2; Gaps 2;
```

DB 1 CTGCCATAGATGGCGTCCGCTGCACGACGAGTTCCTGTGGGCGAGTGGCGTAC 60


```

OY 1330 CGTTAAGACATGTCAGTGGCCCTTTGGAAAAGTGTGGGATCCAGTTCAGAGGA 1389
    |||||||
DB 1022 CGTTAAGACATGTCAGTGGCCCTTTGGAAAAGTGTGGGATCCAGTTCAGAGGA 1081
OY 1390 GGGGATGTTGTCTCCAGTGTGGGCAAGAGTGTATGCGGAATCTCTGCAATAGA 1449
    |||||||
DB 1082 GGGGATGTTGTCTCCAGTGTGGGCAAGAGTGTATGCGGAATCTCTGCAATAGA 1141
OY 1450 GAAAAGCTCCACCATTTCTTTGATGTGCTTTAAATCTCAGCTTCTATATAGA 1509
    |||||||
DB 1142 GAAAAGCTCCACCATTTCTTTGATGTGCTTTAAATCTCAGCTTCTATATAGA 1201
OY 1510 AACAGAGCTCTGTGAGCTCCTGTGTGCTGATGTGTGGAATGATGATGATCAG 1569
    |||||||
DB 1202 AACAGAGCTCTGTGAGCTCCTGTGTGCTGATGTGTGGAATGATGATGATCAG 1261
OY 1570 AAAGCATTTTCTTTCTTTGAACTTAAAGTCTTATTTATTAAGACAGCAGATTC 1629
    |||||||
DB 1262 AAAGCATTTTCTTTCTTTGAACTTAAAGTCTTATTTATTAAGACAGCAGATTC 1321
OY 1630 ACATTTTATACATGAGATCTTCTTGTGTGATACAGAGATGATGATGATGATGAT 1689
    |||||||
DB 1322 ACATTTTATACATGAGATCTTCTTGTGTGATACAGAGATGATGATGATGATGAT 1380
OY 1690 AAAGAGTTTATGTCCTGAGCTGTGCTAAATTTATCTAATTTCCAGATGCTTTGTAG 1749
    |||||||
DB 1381 AAAGAGTTTATGTCCTGAGCTGTGCTAAATTTATCTAATTTCCAGATGCTTTGTAG 1440
OY 1750 ATGACTGAAATTTGTGTGAGCCACATATTGGAGTCTGATGATTTGATGATGAGGA 1809
    |||||||
DB 1441 ATGACTGAAATTTGTGTGAGCCACATATTGGAGTCTGATGATTTGATGATGAGGA 1500
OY 1810 AAGGCAATCTCATGAGATGATGATTAAGAACCAATCTGCTGCAATCTCAGATTCACAGAA 1869
    |||||||
DB 1501 AAGGCAATCTCATGAGATGATGATTAAGAACCAATCTGCTGCAATCTCAGATTCACAGAA 1560
OY 1870 AAGAGGATTCAGACTGAGAGAGTGTGACATAGAGTGAAGACCAAGACTTTGAAT 1929
    |||||||
DB 1561 AAGAGGATTCAGACTGAGAGAGTGTGACATAGAGTGAAGACCAAGACTTTGAAT 1620
OY 1930 TTGCAAGCTGCTCATGTGTGATTAATTAATCACTGCTGCTTTCTATTTAGTACAAATCT 1989
    |||||||
DB 1621 TTGCAAGCTGCTCATGTGTGATTAATTAATCACTGCTGCTTTCTATTTAGTACAAATCT 1680
OY 1990 AATTTTATGAGTTTAAATTAAGAAAATTTACAA 2028
    |||||||
DB 1681 AATTTTATGAGTTTAAATTAAGAAAATTTACAA 1719

RESULT 9
US-60-453-050-274
: Sequence 274, Application US/60453050
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: LUKE, May
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001457
: CURRENT APPLICATION NUMBER: US/60/453,050
: CURRENT FILING DATE: 2003-03-10
: NUMBER OF SEQ ID NOS: 82762
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 274
: LENGTH: 1719
: TYPE: DNA
: ORGANISM: Homo sapiens
US-60-453-050-274

Query Match 66.3% Score 1353.8 DB 11: Length 1719;
Best Local Similarity 94.3% Pred. No. 0;
Matches 1451; Conservative 6; Mismatches 3; Indels 79; Gaps 2;
OY 568 CAGAGCTTGCAGTTTGAAGAAACAGAAATCTCTATATGATTTCTCCATCGTCA 627

```

```

DB 182 CAGAGCTTGCAGTTTGAAGAAACAGAAATCTCTATATGATTTCTCATCTGCA 241
    |||||||
OY 628 ACATACGAGAGAGAGTGTGTGTCCTTGGAGTCTCTGAAATCTTCAACCTTGCAT 687
    |||||||
DB 242 ACATACGAGAGAGAGTGTGTGTCCTTGGAGTCTCTGAAATCTTCAACCTTGCAT 301
OY 688 CCTTGGCTTCATGCTGACAGCACTTCTAGATGAGCCCTTGGATTAAGAGGTGATCA 747
    |||||||
DB 302 CCTTGGCTTCATGCTGACAGCACTTCTAGATGAGCCCTTGGATTAAGAGGTGATCA 361
OY 748 CATGCTGTGCTGTGAGCTTGTGAGAGAACTAAGAGATTTTCACAGAGCTTATATAGAA 807
    |||||||
DB 362 CATGCTGTGCTGTGAGCTTGTGAGAGAACTAAGAGATTTTCACAGAGCTTATATAGAA 421
OY 808 GAATGTTTGGAGAAATATGACAGAGTATCAACATGATCTTCAATCAATTTGGCTTCAA 867
    |||||||
DB 422 GAATGTTTGGAGAAATATGACAGAGTATCAACATGATCTTCAATCAATTTGGCTTCAA 481
OY 868 GGAATTTACAGAGTACCTGATCACTGAGGAAATGACACAGAGAGTATGATCAACAGCT 927
    |||||||
DB 482 GGAATTTACAGAGTACCTGATCACTGAGGAAATGACACAGAGAGTATGATCAACAGCT 541
OY 928 TCTAAGAAAG----- 938
    |||||||
DB 542 TCTAAGAAAGATTTGAGGCTGTGAACAGAACTAAGATATGCGCGAAACAAA 601
OY 939 -----GACCTGCTCCCATTTGTCCTCCCTGTCTATAG 969
    |||||||
DB 602 CCGATGGCTTAAACACGCTTTTGTGAGCAGACCTGTGCTCCATTTGTCCTCCCTGTCTATAG 661
OY 970 CTTAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029
    |||||||
DB 662 CTTAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 721
OY 1030 CGTGAAGTTTCAATCAGAGGCGCACAGGCTTACAGGCTTCAACATTAAGATCCATACAA 1089
    |||||||
DB 722 TGTGCAAAATTTTCATCCAGGCGCACAGGCTTACAGGCTTCAACATTAAGATCCATACAA 781
OY 1090 TGAAGCTGAGAACACAGAGATTTATCACCCTGTGTGATCTGTGATGATGATGATGATGAT 1149
    |||||||
DB 782 TGAAGCTGAGAACACAGAGATTTATCACCCTGTGTGATCTGTGATGATGATGATGATGAT 841
OY 1150 GGAATGCGAATGGGCGAGGCGACATTAATCCAAATCCCACTGAACCACTGAAGAAAG 1209
    |||||||
DB 842 GGAATGCGAATGGGCGAGGCGACATTAATCCAAATCCCACTGAACCACTGAAGAAAG 901
OY 1210 AAGAGATTCAGACTGAGATGCTGTCAACACATAGAAATGATGATGATGATGATGATGATGAT 1269
    |||||||
DB 902 AAGAGATTCAGACTGAGATGCTGTCAACACATAGAAATGATGATGATGATGATGATGATGAT 961
OY 1270 TAACAAAGAACTTAAAGGAGAGATCCCGAGGCGAGAAATGATCAAGAGCTGAATGCGAG 1329
    |||||||
DB 962 TAACAAAGAACTTAAAGGAGAGATCCCGAGGCGAGAAATGATCAAGAGCTGAATGCGAG 1021
OY 1330 CGTTAAGAGACATGTCAGTGGCCCTTTGGAAAAGTGTGGGATCCAGTTCAGAGGA 1389
    |||||||
DB 1022 CGTTAAGAGACATGTCAGTGGCCCTTTGGAAAAGTGTGGGATCCAGTTCAGAGGA 1081
OY 1390 GGGGATGTTGTCTCCAGTGTGGGCAAGAGTGTATGCGGAATCTCTGCAATAGA 1449
    |||||||
DB 1082 GGGGATGTTGTCTCCAGTGTGGGCAAGAGTGTATGCGGAATCTCTGCAATAGA 1141
OY 1450 GAAAAGCTCCACCATTTCTTTGATGTGCTTTAAATCTCAGCTTCTATATAGA 1509
    |||||||
DB 1142 GAAAAGCTCCACCATTTCTTTGATGTGCTTTAAATCTCAGCTTCTATATAGA 1201
OY 1510 AACAGAGCTCTGTGAGCTCCTGTGTGCTGATGTGTGGAATGATGATGATGATGATGAT 1569
    |||||||
DB 1202 AACAGAGCTCTGTGAGCTCCTGTGTGCTGATGTGTGGAATGATGATGATGATGATGAT 1261
OY 1570 AAAGCATTTTCTTTCTTTGAACTTAAAGTCTTATTTATTAAGACAGCAGATTC 1629
    |||||||

```


Db 722 CGTCAAGTTTCATCCAGGGCCCAAGCCTACAGCCATCCATAAAGATGCCATACAA 781
 Oy 1090 TGAAGCTGAGAACAGAGAGATTATACCTGTGTGACCTCTGTGATCGAATCATCATGG 1149
 Db 782 TGAAGCTGAGAACAGAGAGATTATACCTGTGTGACCTCTGTGATCGAATCATCATGG 841
 Oy 1150 GATTCGCCGAATGGGCGACCCACTATAATCCCAATCCCACTTAACCACTGAAGAAAG 1209
 Db 842 GGATCCCGAATGGGCGACCCACTATAATCCCACTTAACCACTGAAGAAAG 901
 Oy 1210 AAGAAGATTGGACTGAGATGCTGTCAACACATAGAAAGTCAAGATGTTTCCCGACCTA 1269
 Db 902 AAGAAGATTGGACTGAGATGCTGTCAACACATAGAAAGTCAAGATGTTTCCCGACCTA 961
 Oy 1270 TAACAAAGAACCTTAAGAGGAGATCCCGAGGCGAGAAATGATCAAGAGCTGAATGCA 1329
 Db 962 TAACAAAGAACCTTAAGAGGAGATCCCGAGGCGAGAAATGATCAAGAGCTGAATGCA 1021
 Oy 1330 CGTTTAAGAGACATGTCCAGTGGCTTTGGAAAGTGTGGGATCCAGTTCAGAGGGA 1389
 Db 1022 CGTTTAAGAGACATGTCCAGTGGCTTTGGAAAGTGTGGGATCCAGTTCAGAGGGA 1081
 Oy 1390 GGGGTATGTTTGTCTCCAGTGTGGGCAAGAGTGTCTATGCGGAATTCCTCATAGCA 1449
 Db 1082 GGGGTATGTTTGTCTCCAGTGTGGGCAAGAGTGTCTATGCGGAATTCCTCATAGCA 1141
 Oy 1450 GAAAGCTCCACCATTTTCTTTGATGTGTTTAAAGTCTACAGTCTCTTAATAGA 1509
 Db 1142 GAAAGCTCCACCATTTTCTTTGATGTGTTTAAAGTCTACAGTCTCTTAATAGA 1201
 Oy 1510 AACAGAGTCTGTGACGCTGTGTGTGGCTGATGTCTGGAATGATGATGCTCAGG 1569
 Db 1202 AACAGAGTCTGTGACGCTGTGTGTGGCTGATGTCTGGAATGATGATGCTCAGG 1261
 Oy 1570 AAGGATTTTCTTTTCTTTGATGATCTTAAGGTTCTATTATTAAGGACACAGATTC 1629
 Db 1262 AAGGATTTTCTTTTCTTTGATGATCTTAAGGTTCTATTATTAAGGACACAGATTC 1321
 Oy 1630 ACATTTTATACATGAGATCTTCTTTGTGTGAATACAGAGATTGACTGCTCATCTT 1689
 Db 1322 ACATTTTATACATGAGATCTTCTTTGTGTGAATACAGAGATTGACTGCTCATCTT 1380
 Oy 1690 AAGAAGTTTATGTCCCTGTGACTGTGGCTAAATATCTAATTTCCAGATGCTTTGTAG 1749
 Db 1381 AAGAAGTTTATGTCCCTGTGACTGTGGCTAAATATCTAATTTCCAGATGCTTTGTAG 1440
 Oy 1750 ATGACGAGATTTGTGAGCCACATATTGGAGTTCTAGATTGAGTGAATGGCAGGA 1809
 Db 1441 ATGACGAGATTTGTGAGCCACATATTGGAGTTCTAGATTGAGTGAATGGCAGGA 1500
 Oy 1810 AGGGCATCTCATTTGATGATTAAGTGAACCAACTAGTCTCGAATTTCTACAGAGA 1869
 Db 1501 AGGGCATCTCATTTGATGATTAAGTGAACCAACTAGTCTCGAATTTCTACAGAGA 1560
 Oy 1870 AGAGAGGATCAGACTGAGAGAGCTGTGACATAGGACTTAAGACCAAGACTTTGAAT 1929
 Db 1561 AGAGAGGATCAGACTGAGAGAGCTGTGACATAGGACTTAAGACCAAGACTTTGAAT 1620
 Oy 1930 TTGCGAGCTCTCATGTGTGAGTTATATACAGTGTCTTTCTATGAGTTCAATCT 1989
 Db 1621 TTGCGAGCTCTCATGTGTGAGTTATATACAGTGTCTTTCTATGAGTTCAATCT 1680
 Oy 1990 ATATTTTATGAGTTTAAATTAAGAAAAATTTACAA 2028
 Db 1681 ATATTTTATGAGTTTAAATTAAGAAAAATTTACAA 1719

RESULT 8
 US-60-453-135-274
 ; Sequence 274, Application US/60453135
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; APPLICANT: IAKOUBOVA, Olgia

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001456
 ; CURRENT APPLICATION NUMBER: US/60/453,135
 ; CURRENT FILING DATE: 2003-03-10
 ; NUMBER OF SEQ ID NOS: 82762
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 274
 ; LENGTH: 1719
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-60-453-135-274
 Query Match 66.3%; Score 1353.8; DB 11; Length 1719;
 Best Local Similarity 94.3%; Pred. No. 0;
 Matches 1451; Conservative 6; Mismatches 3; Indels 79; Gaps 2;
 Oy 568 CAGAGCTTGCAGACTTTTGAAGAAACAGAAATCTCTCATAGTAATTTCTCATGCTCA 627
 Db 182 CAGAGCTTGCAGAGTTTGAAGAAACAGAAATCTCTCATAGTAATTTCTCATGCTCA 241
 Oy 628 ACATAGGGAAGAGTGTGTGCTCCCTTGGAGGCTCTGAGTCTTAACCTTGAT 687
 Db 242 ACATAGGGAAGAGTGTGTGCTCCCTTGGAGGCTCTTAAGTCTTAACCTTGAT 301
 Oy 688 CCTTTGGCTTCATGCTGACAGGACAGTTCTAGATGAGCGCTGGATTAAGAGGTGATGA 747
 Db 302 CCTTTGGCTTCATGCTGACAGGACAGTTCTAGATGAGCGCTGGATTAAGAGGTGATGA 361
 Oy 748 CATGCTTGTGCTGTGGGCTCTTGGAGAGACTAAGAGATTTTCAACAGCCTAATACGAA 807
 Db 362 CATGCTTGTGCTGTGGGCTCTTGGAGAGACTAAGAGATTTTCAACAGCCTAATACGAA 421
 Oy 808 GAATGTTTGGAAATTAAGCCAGGACTATCAATGATATCTTCAATCAATGGCTTCAA 867
 Db 422 GAATGTTTGGAAATTAAGCCAGGACTATCAATGATATCTTCAATCAATGGCTTCAA 481
 Oy 868 GGAATTTTCAGAGTACCTGATCACTGAGGAGAAATGACACATGAGACTAGTAACAGCT 927
 Db 482 GGAATTTTCAGAGTACCTGATCACTGAGGAGAAATGACACATGAGACTAGTAACAGCT 541
 Oy 928 TCTTAAGAAAG----- 938
 Db 542 TCTTAAGAAAGTATTGAGGCTCTGAACAAGTAAGATATGCCCCGAAACAAA 601
 Oy 939 -----GACCTGGTCCCATTTGTCCTCCCTGCTATG 969
 Db 602 CCGATGGGTTAAAAACCGTTTTTTGAGCAGACCTGGTCCCATTTGCCCCCTGTATGG 661
 Oy 970 CTTAGAGATTCATGATGCTCGAAGTGGAGAGTGTGTTCTTGAACCTGCTTGAAT 1029
 Db 662 CTTAGAGATTCATGATGCTCGAAGTGGAGAGTGTGTTCTTGAACCTGCTTGAAT 721
 Oy 1030 CGTGCAAGTTTATCCAGGGCCACAGCCCTACAGCCACATCAATAAAGATGCCATCAA 1089
 Db 722 YGTGCAAGTTTATCCAGGGCCACAGCCCTACAGCCACATCAATAAAGATGCCATCAA 781
 Oy 1090 TGAAGCTGAGAACAGAGAGATTATACCTGTGTGACCTCTGTGATGGAATCATATGG 1149
 Db 782 TGAAGCTGAGAACAGAGAGATTATACCTGTGTGACCTCTGTGATGGAATCATATGG 841
 Oy 1150 GGATCCCGAATGGGCGACCCACTATAATCCCAATCCCACTTAACCACTGAAGAAAG 1209
 Db 842 GGATCCCGAATGGGCGACCCACTATAATCCCAATCCCACTTAACCACTGAAGAAAG 901
 Oy 1210 AAGAAGATTGGACTGAGATGCTGTCAACACATAGAAAGTCAAGATGTTTCCCGACCTA 1269
 Db 902 AAGAAGATTGGACTGAGATGCTGTCAACACATAGAAAGTCAAGATGTTTCCCGACCA 961
 Oy 1270 TAACAAAGAACCTTAAGAGGAGATCCCGAGGCGAGAAATGATCAAGAGCTGAATGCA 1329
 Db 962 TAACAAAGAACCTTAAGAGGAGATCCCGAGGCGAGAAATGATCAAGAGCTGAATGCA 1021

Db 481 TCAATTGGCTTCAAGAAATTTACAGATACCTGATCATCTGAGGAAATATGACACTGAG 540
 QY 914 ACTAGTACCACTCTTAAAGAAAG----- 939
 Db 541 ACTAGTACCACTCTTAAAGAAAGTATGAGCTCTGAAACAGTAACTAGAGATAT 600
 QY 940 -----ACCTGGTCCCATTTGTC 955
 Db 601 GCCCGGAAACAAACGATGGGTTAAAAACGGTTTTTTGAGCAGACCTGGTCCCATTTGTC 660
 QY 956 CCCCCCTGTATGCTTAGAGATATCATGATGTCGAAAGTGGAGAGAGATGTTCTTGA 1015
 Db 661 CCCCCCTGTATGCTTAGAGATATCATGATGTCGAAAGTGGAGAGATGTTCTTGA 720
 QY 1016 CCTGCTCTTGAATTCGTGCAAAAGTTTCATCCAGGGCCACAAAGCTTACAGCCATCCATA 1075
 Db 721 CCTGCTCTTGAATTCGTGCAAAAGTTTCATCCAGGGCCACAAAGCTTACAGCCATCCATA 780
 QY 1076 AAGATGCATACATGATGAGCTGAGAACAGAGAGATATGATCAGCTGTGACCTGTGAT 1135
 Db 781 AAGATGCATACATGATGAGCTGAGAACAGAGAGATATGATCAGCTGTGACCTGTGAT 840
 QY 1136 CGAATCATCTATGGGATCGCGAATGGCGACGACATTAATCCAAATCCCACTTGAAC 1195
 Db 841 CGAATCATCTATGGGATCGCGAATGGCGACGACATTAATCCAAATCCCACTTGAAC 900
 QY 1196 CAACCTGAAGAAAGAAATTTGAGTCTGATCTGCTCAACACCATAGAAAGCAAGT 1255
 Db 901 CAACCTGAAGAAAGAAATTTGAGTCTGATCTGCTCAACACCATAGAAAGCAAGT 960
 QY 1256 GTTTCCTCCAGACTATACAAAGAACTTAAAGGAGGATCCCGGAGGAGCAATGATCAA 1315
 Db 961 GTTTCCTCCAGACTATACAAAGAACTTAAAGGAGGATCCCGGAGGAGCAATGATCAA 1020
 QY 1316 GAGCTGAATGCAGCGTTTAAAGACATGTCAGATGCGCTTTGGAAGAGTGGTGGATC 1375
 Db 1021 GAGCTGAATGCAGCGTTTAAAGACATGTCAGATGCGCTTTGGAAGAGTGGTGGATC 1080
 QY 1376 CAGTTCAAGGAGGAGGATGTTTGTCTCCCACTCTGGGCAAGAGATGCTATGCCGA 1435
 Db 1081 CAGTTCAAGGAGGAGGATGTTTGTCTCCCACTCTGGGCAAGAGATGCTATGCCGA 1140
 QY 1436 TTCTCTCATAGCAGAAAGCTCCACCATTTTCTTTGATGATGTTTAAAGCTCAGC 1495
 Db 1141 TTCTCTCATAGCAGAAAGCTCCACCATTTTCTTTGATGATGTTTAAAGCTCAGC 1200
 QY 1496 TTCTCTATAATAGAAACAGAGGCTGTGACGCTCTGTGTGGCTGATGTCTGAAA 1555
 Db 1201 TTCTCTATAATAGAAACAGAGGCTGTGACGCTCTGTGTGGCTGATGTCTGAAA 1260
 QY 1556 TGATGTAGTTCAGGAAAGCAATTTTCTTTGAAACCTTAAAGTTCTATTTAA 1615
 Db 1261 TGATGTAGTTCAGGAAAGCAATTTTCTTTGAAACCTTAAAGTTCTATTTAA 1320
 QY 1616 GCAGCAAGATTCACATTTTATACATGAGGATCTTCTGTGTGAATCCCGGATG 1675
 Db 1321 GCAGCAAGATTCACATTTTATACATGAGGATCTTCTGTGTGAATCCCGGATG 1380
 QY 1676 ACTGCATCCCTTTAAAGAAAGTTTATGCTCCCTGACTCTGGCTAAATTTATCTAATTTCC 1735
 Db 1381 ACTGCATCCCTTTTAAAGAAAGTTTATGCTCCCTGACTCTGGCTAAATTTATCTAATTTCC 1439
 QY 1736 AGATGCTTTGTAGATGATGAGTATTTGTGAGCCACATTTGGAGTTCTAGATTTGA 1795
 Db 1440 AGATGCTTTGTAGATGATGAGTATTTGTGAGCCACATTTGGAGTTCTAGATTTGA 1499
 QY 1796 GTGATGAGCAAGAAAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAATTAATTTCCG 1855
 Db 1500 GTGATGAGCAAGAAAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAATTAATTTCCG 1559
 QY 1856 GAATTTACAGAGAGAGGAGGATCACTGAGAGAGCTGTGACATGAGACTTGAAGACC 1915
 Db 1560 GAATTTACAGAGAGAGGAGGATCACTGAGAGAGCTGTGACATGAGACTTGAAGACC 1619

QY 1916 AAGACTTGAATTTGCGAGCTGCTCATGCTGAGTATATACAGTCTCTCTAT 1975
 Db 1620 AAGACTTGAATTTGCGAGCTGCTCATGCTGAGTATATACAGTCTCTCTAT 1679
 QY 1976 TGAGTTCAAAATCTATTTTATTTGAGAGTTTAAATTAAGAAAAATTTACAGAAAAA 2035
 Db 1680 TGAGTTCAAAATCTATTTTATTTGAGAGTTTAAATTAAGAAAAATTTTAAAAA 1739
 QY 2036 AAAAAA 2041
 Db 1740 AAAAAA 1745
 RESULT 7
 US-10-170-235-41123
 ; Sequence 41123, Application US/10170235
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig
 ; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF
 ; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREO
 ; FILE REFERENCE: CL001380
 ; CURRENT APPLICATION NUMBER: US/10/170, 235
 ; CURRENT FILING DATE: 2003-03-17
 ; NUMBER OF SEQ ID NOS: 42514
 ; SEQ ID NO 41123
 ; LENGTH: 1719
 ; TYPE: DNA
 ; ORGANISM: HUMAN
 US-10-170-235-41123
 Query Match 66.4%; Score 1354.6; DB 8; Length 1719;
 Best Local Similarity 94.6%; Pred. No. 0;
 Matches 1456; Conservative 0; Mismatches 4; Indels 79; Gaps 2;
 QY 568 CAGAGCTTGCAGATTTTGAAGAAAGCAATCTCATAGTGAATTTCTCATGCTCA 627
 Db 182 CAGAGCTTGCAGATTTTGAAGAAAGCAATCTCATAGTGAATTTCTCATGCTCA 241
 QY 628 ACATACGAGAGAGAGTGTGCTCCCTTGGAGGCTCTCTGAATCTCTTAACCTTGCAT 687
 Db 242 ACATACGAGAGAGAGTGTGCTCCCTTGGAGGCTCTCTGAATCTCTTAACCTTGCAT 301
 QY 688 CCTTGGCTCATGCTGACCGAGGAGTCTCATAGTGAAGGCTTGAAGAGGAGTGA 747
 Db 302 CCTTGGCTCATGCTGACCGAGGAGTCTCATAGTGAAGGCTTGAAGAGGAGTGA 361
 QY 748 CATGCTTGTGCTGGGCTCTTGGAGAACTTAAGATTTTCAAGAGCTTAATATCAGAA 807
 Db 362 CATGCTTGTGCTGGGCTCTTGGAGAACTTAAGATTTTCAAGAGCTTAATATCAGAA 421
 QY 808 GAATGTTTGGAAATTAAGCAGAGCTATCAACATGATCTTCCATCAATTTGGCTTCAA 867
 Db 422 GAATGTTTGGAAATTAAGCAGAGCTATCAACATGATCTTCCATCAATTTGGCTTCAA 481
 QY 868 GGAATTTCAAGAGTACCTGATCACTGAGGAGAAATGACACTGGAGACTGATACAGCT 927
 Db 482 GGAATTTCAAGAGTACCTGATCACTGAGGAGAAATGACACTGGAGACTGATACAGCT 541
 QY 928 TCTAAGAAAG----- 938
 Db 542 TCTAAGAAAGTATGAGGCTCTGAACAAGTAAAGATATGCCGGAAACAAA 601
 QY 939 -----GACCTGTGCCATTTGTCCTCCCTGTCTATGG 969
 Db 602 CCGATGGTTTAAACCGTTTGTGAGCAGACCTGTGCCATTTGCCCTGTCTATGG 661
 QY 970 CTTAGAGTATCTATGTTCTGAAGTGGGAGAGTCTGTCTTAACCTGCTCTGAAT 1029
 Db 662 CTTAGAGTATCTATGTTCTGAAGTGGGAGAGTCTGTCTTAACCTGCTCTGAAT 721
 QY 1030 CGTCAAGAGTTTCTCAAGGGCCCAAGCCCTACAGCCACTGCATTAAGATGCCATCAA 1089

Db 961 GCTGTAAACAGTAAGTAAGATATGCGGAAACAAACCGATGGTTAAAAACGT 1020
QY 939 -----GACCTGGTCCATTGTCCCTGCTCTATGCTTAGAGTATCTGATGTC 988
Db 1021 TTTTGGACGACGCTGGTCCATGTGCCCTGTCTAAGGCTTAGAGTATCTGATGTC 1080
QY 989 TCGAAGTGGAGAGTATCTTCTTAACCTGCTCTTGAATAGTGCAGAAAGTTTCATCCAG 1048
Db 1081 TCGAAGTGGAGAGTATCTTCTTAACCTGCTCTTGAATAGTGCAGAAAGTTTCATCCAG 1140
QY 1049 GCGCACAGCCCTACAGCCACTCCCAATTAAGATGCCATACATGAAGCTGAGAACAGAGA 1108
Db 1141 GCGCACAGCCCTACAGCCACTCCCAATTAAGATGCCATACATGAAGCTGAGAACAGAGA 1200
QY 1109 AGTTATCAGCTGTGAGCTCTGTGATGATGATCATCATTTGGGATGCGCAATGGCAGCG 1168
Db 1201 AGTTATCAGCTGTGAGCTCTGTGATGATGATCATCATTTGGGATGCGCAATGGCAGCG 1260
QY 1169 CACATTAATCCAAATCCCACTTGAACCACTGAAGAAAGAAAGAAAGATGGACTCAGAT 1228
Db 1261 CACATTAATCCAAATCCCACTTGAACCACTGAAGAAAGAAAGAAAGATGGACTCAGAT 1320
QY 1229 GCTGTCAACACCATAGAAAGTACAGAGTGTTCCTCCAGACTATACAGAAACCTAAAGG 1288
Db 1321 GCTGTCAACACCATAGAAAGTACAGAGTGTTCCTCCAGACTATACAGAAACCTAAAGG 1380
QY 1289 AAGGATCCCAAGGCGACAGATGATGAAGAGCTGAATGACGCGTTTAAGAGACATGTCCA 1348
Db 1381 AAGGATCCCAAGGCGACAGATGATGAAGAGCTGAATGACGCGTTTAAGAGACATGTCCA 1440
QY 1349 GTGGCTTTGAAAGGTGGTGGGATCCAGTTACAGAGGAGGAGGTATGTTCTCCCA 1408
Db 1441 GTGGCTTTGAAAGGTGGTGGGATCCAGTTACAGAGGAGGAGGTATGTTCTCCCA 1500
QY 1409 GTCTGGGCAAGAGAGTGTATGCGGAATTCCTGCATAGCAGAAAGCTCCACCATTTT 1468
Db 1501 GTCTGGGCAAGAGAGTGTATGCGGAATTCCTGCATAGCAGAAAGCTCCACCATTTT 1560
QY 1469 CTTTGTGATGCTTTTAAAGTCTCAGTCTCTATATAGAAACAGCAGGCTTGTCCAGC 1528
Db 1561 CTTTGTGATGCTTTTAAAGTCTCAGTCTCTATATAGAAACAGCAGGCTTGTCCAGC 1620
QY 1529 TCCCTGTGCTGATGCTGCTGGAATGATGATGATGATGATGATGATGATGATGATGAT 1588
Db 1621 TCCCTGTGCTGATGCTGCTGGAATGATGATGATGATGATGATGATGATGATGATGAT 1680
QY 1589 TGAACCTTAAGGTTCTATTTATTAAGCAGACAGATTCACATTTTATACATGAGGA 1648
Db 1681 TGAACCTTAAGGTTCTATTTATTAAGCAGACAGATTCACATTTTATACATGAGGA 1740
QY 1649 TCTTCTTTGTGGTAATACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1708
Db 1741 TCTTCTTTGTGGTAATACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1799
QY 1709 GACCTGTGCTAAATATATCTAATTTCCAGATGCTTTTGTAGATGCTAAGTATTTGGA 1768
Db 1800 GACCTGTGCTAAATATATCTAATTTCCAGATGCTTTTGTAGATGCTAAGTATTTGGA 1859
QY 1769 GCCACATTTTGGAGTTTAAATTTGAGTGAATGGCAGAAAGGCGCATTCOCATTTGGA 1828
Db 1860 GCCACATTTTGGAGTTTAAATTTGAGTGAATGGCAGAAAGGCGCATTCOCATTTGGA 1919
QY 1829 TGATTAAGTGAACCAACTAGTCTCGGAATTTACAGAGAAGGAGGAGAAATCAGACTAG 1888
Db 1920 TGATTAAGTGAACCAACTAGTCTCGGAATTTACAGAGAAGGAGGAGAAATCAGACTAG 1979
QY 1889 GAACCTGTGACATGAGCTTAAGACCAAGACATTTGAATTTGGAGCTCTCATGTGT 1948
Db 1980 GAACCTGTGACATGAGCTTAAGACCAAGACATTTGAATTTGGAGCTCTCATGTGT 2039
QY 1949 GAGTATATATCATGCTCTCTTCTATGATGATGATGATGATGATGATGATGATGATGAT 2008
Db 2040 GAGTATATATCATGCTCTCTTCTATGATGATGATGATGATGATGATGATGATGATGAT 2099

QY 2009 AATTAAGAAAAATTTACAA 2028
Db 2100 AATTAAGAAAAATTTACAA 2119
RESULT 6
US-10-342-887-1692
; Sequence 1692, Application US/10342887
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT FILING DATE: 2003-01-15
; PRIOR FILING DATE: 2003-01-15
; PRIOR FILING DATE: 2001-06-18
; PRIOR FILING DATE: 2002-05-14
; PRIOR FILING DATE: 2002-05-14
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1692
; LENGTH: 1749
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1692
Query Match 76.3%; Score 1556.8; DB 9; Length 1749;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 7; Indels 79; Gaps 2;
QY 374 ACCAATTTATTAATGAATCTGCTGCGGAAGTCTTGTCAATACCAAGCCAGAGG 433
Db 1 ACCAATTTATTAATGAATCTGCTGCGGAAGTCTTGTCAATACCAAGCCAGAGG 60
QY 434 ATGGCACTGAGAAAGTGAATGACCGAAAGTGAGCTTGAAGAGAGATGCTTGTGA 493
Db 61 ATGGCACTGAGAAAGTGAATGACCGAAAGTGAGCTTGAAGAGAGATGCTTGTGA 120
QY 494 CTTCACAAAGCCTTAAGCAGAGTGAACCCAGAAATGCTGCCAAGCTGCATCAGATGAC 553
Db 121 CTTCACAAAGCCTTAAGCAGAGTGAACCCAGAAATGCTGCCAAGCTGCATCAGATGAC 180
QY 554 AAAGCAAGAGTGGCGAGAGGCTTGAAGTTTGAAGAAAGCAAGATCTCATAGTGA 613
Db 181 AAAGCAAGAGTGGCGAGAGGCTTGAAGTTTGAAGAAAGCAAGATCTCATAGTGA 240
QY 614 TTTCTCATGCTCAACATACGGAAGAGTGGTCCCTTGAGAGTCTCTGAAGTTC 673
Db 241 TTTCTCATGCTCAACATACGGAAGAGTGGTCCCTTGAGAGTCTCTGAAGTTC 300
QY 674 TCTAACCTTGCATCTTTGGCTTCAATGCTGACAGCAGAGTTCTAGATGAGCGTTGAT 733
Db 301 TCTAACCTTGCATCTTTGGCTTCAATGCTGACAGCAGAGTTCTAGATGAGCGTTGAT 360
QY 734 AAGAGGAGTGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
Db 361 AAGAGGAGTGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 794 CGCTATATCGAAGAAATGTTTGGGAAATAGCCAGAGCTATCAACATGATGATCTTCCAA 853
Db 421 CGCTATATCGAAGAAATGTTTGGGAAATAGCCAGAGCTATCAACATGATGATCTTCCAA 480
QY 854 TCAATGCTTCAGAGAAATTCAGAGTACCTGATGATGATGATGATGATGATGATGATGATGAT 913

Query Match	91.6%	Score 1869.6	DB 11	Length 2119
Best Local Similarity	94.4%	Pred. No. 0		
Matches 2002	Conservative	8	Mismatches 7	Indels 103
			Gaps	3
QY 11	ATGGCGTCCGTCGGGGCGCTGCACAGACATCTCTGTGGGCACTGGGCTTAGAGGGCGCTCGAA	70		
Db 1	ATGGCGTCCGTCGGGGCGCTGCACAGACATCTCTGTGGGCACTGGGCTTAGAGGGCGCTCGAA	60		
QY 71	CGGAGCCCTACCTCTTGTATGATTCCTCGGGGCGACGGGACCGGGAAATTCACGCTGGCG	130		
Db 61	CGGAGCCCTACCTCTTGTATGATTCCTCGGGGCGACGGGACCGGGAAATTCACGCTGGCG	120		
QY 131	TTGGAGCTAGGCGCCAGCGGCTCGGGGTAGATTCGTACAGCGCTGACTCCATCGAGGCTCTAT	190		
Db 121	TTGGAGCTAGGCGCCAGCGGCTCGGGGTAGATTCGTACAGCGCTGACTCCATCGAGGCTCTAT	180		
QY 191	GAAGGCTTAGACATCATACCAACAAGGTTCTGCCCCAAGACAGAGAAATCTCCCGCAC	250		
Db 181	GAAGGCTTAGACATCATACCAACAAGGTTCTGCCCCAAGACAGAGAAATCTCCCGCAC	240		
QY 251	CACATGATCAGCTTGTGTGATCTCTGTGACCAATTCACAGTGTGGAGCTTCGAAT	310		
Db 241	CACATGATCAGCTTGTGTGATCTCTGTGACCAATTCACAGTGTGGAGCTTCGAAT	300		
QY 311	AGAGCACTGCTGTGATGAA-----GATATATTTGCCCGA	346		
Db 301	AGAGCACTGCTGTGATATTCCTTAGAANAAGCAGCAGCTGGTGTGATATATTTGCCCGA	360		
QY 347	GACAAATTCCTATTTGTGTGGAGGAACCAATTTATACATTAATCTCTGCTGGAAA	406		
Db 361	GACAAATTCCTATTTGTGTGGAGGAACCAATTTATACATTAATCTCTGCTGGAAA	420		
QY 467	GTTCTGTCAATACCAACCCCGAGAGATGGGCACTGAGAAATGATGACCGAAGAG	466		
Db 421	GTTCTGTCAATACCAACCCCGAGAGATGGGCACTGAGAAATGATGATGACCGAAGAG	480		
QY 467	GAGCTTGAAGAGAGATGTCTTGTACTTCAACAACCCCTAAGCCAGGTGGACCCAGAA	526		
Db 481	GAGCTTGAAGAGAGATGTCTTGTACTTCAACAACCCCTAAGCCAGGTGGACCCAGAA	540		
QY 527	ATGGCTGCCAAGCTGCATCCACATGACAAACGCAAGTGGCCGAGAGCTTGCAGATTTT	586		
Db 541	ATGGCTGCCAAGCTGCATCCACATGACAAACGCAAGTGGCCGAGAGCTTGCAGATTTT	600		
QY 587	GAGGAAACAGGAATCTTCATAGTAATTTCTCCATCGTCAACATACGGAAGAAGTGT	646		
Db 601	GAGGAAACAGGAATCTTCATAGTAATTTCTCCATCGTCAACATACGGAAGAAGTGT	660		
QY 647	GSTGCCCTTGGAGTCTCTGTGAAGTTCTTACCCCTTGATCCTTTGGCTTATGCTGAC	706		
Db 661	GSTGCCCTTGGAGTCTCTGTGAAGTTCTTACCCCTTGATCCTTTGGCTTATGCTGAC	720		
QY 707	CAGGCACTTCTAGATGAGCGCTTGGATAAGAGAGGTGATGACATGCTTGCTCTGGGCTC	766		
Db 721	CAGGCACTTCTAGATGAGCGCTTGGATAAGAGAGGTGATGACATGCTTGCTCTGGGCTC	780		
QY 767	TTGGAGGAACCTTAGAGATTTTACACAGAGGCTTATAATCAGAGAAATGTTCCGAAAATAC	826		
Db 781	TTGGAGGAACCTTAGAGATTTTACACAGAGGCTTATAATCAGAGAAATGTTCCGAAAATAC	840		
QY 827	CAGGCACTTCAATCATGATCTTCCAAATCAATTTGGCTTTCAGAAATTTACAGGATACCTG	886		
Db 841	CAGGCACTTCAATCATGATCTTCCAAATCAATTTGGCTTTCAGAAATTTACAGGATACCTG	900		
QY 887	ATCACTGAGGGAATAATGCACACTGAGAGCTAGATAACAGCTTCTTAAAGAAAG-----	938		
Db 901	ATCACTGAGGGAATAATGCACACTGAGAGCTAGATAACAGCTTCTTAAAGAAAGATTTGAG	960		
QY 939	-----	938		

```
Db 1621 TCCTTTGTGGCTGATGTCTGGAATGATGATTCAGGAAAGCATTTTCTT 1680
Oy 1589 TGAACCTTAAGCTTCTATATATTAAGAGCAGACAGATTCACATTTTATACATGAGA 1648
Db 1681 TGAACCTTAAGCTTCTATATATTAAGAGCAGACAGATTCACATTTTATACATGAGA 1740
Oy 1649 TCTTCTTTGTGGTGAATACAGAGATTGACTGCATCCCTTTAAAGAGTTTATGTCCT 1708
Db 1741 TCTTCTTTGTGGTGAATACAGAGATTGACTGCATCCCTTTAAAGAGTTTATGTCCT 1799
Oy 1709 GACTGTGGCTAAATATATCAATTTCCAGATGCTTTGTGATGACTGAGATTTGTGA 1768
Db 1800 GACTGTGGCTAAATATATCAATTTCCAGATGCTTTGTGATGACTGAGATTTGTGA 1859
Oy 1769 GCCACATATTGGAGTTCTAGATTTGAGTAATGGCAGGAAGGCCATCTCCATTGAGA 1828
Db 1860 GCCACATATTGGAGTTCTAGATTTGAGTAATGGCAGGAAGGCCATCTCCATTGAGA 1919
Oy 1829 TGATTAAGTGAACCAACTGTTCTCGGATTTCTACAGAAAGAGGAAATCAGACTGAG 1888
Db 1920 TGATTAAGTGAACCAACTGTTCTCGGATTTCTACAGAAAGAGGAAATCAGACTGAG 1979
Oy 1889 GAAGCTGTGACATAGGACTGGAAGACCAAGACTTTGAATTTGCGAGCTGCTCATGTGT 1948
Db 1980 GAAGCTGTGACATAGGACTGGAAGACCAAGACTTTGAATTTGCGAGCTGCTCATGTGT 2039
Oy 1949 GACTTATATACAGTGTCTTCTATATTAAGTACAAATCTAATTTTATTAAGTTTA 2008
Db 2040 GACTTATATACAGTGTCTTCTATATTAAGTACAAATCTAATTTTATTAAGTTTA 2099
Oy 2009 AATAAGAAAAAATTTACAA 2028
Db 2100 AATAAGAAAAAATTTACAA 2119

RESULT 4
US-60-453-135-275
; Sequence 275, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001436
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 275
; LENGTH: 2119
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-275

Query Match 91.6%; Score 1869.6; DB 11; Length 2119;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 2002; Conservative 8; Mismatches 7; Indels 103; Gaps 3;
```

```
Oy 251 CACATGATCAGCTTTGTGATCTCTTGTGACCAATTAACAGTGTGGACTTGAAT 310
Db 241 CACATGATCAGCTTTGTGATCTCTTGTGACCAATTAACAGTGTGGACTTGAAT 300
Oy 311 AGAGCAACTGCTCGATTGAA-----GATATATTGGCCGA 346
Db 301 AGAGCAACTGCTCGATTGAA-----GATATATTGGCCGA 360
Oy 347 GACAAATTCCTATTGTTGTGGAGAACCAATTAATTAATTAATTCCTGCTGAGAA 406
Db 361 GACAAATTCCTATTGTTGTGGAGAACCAATTAATTAATTAATTCCTGCTGAGAA 420
Oy 407 GTTCTTGTCAATACCAAGCCCGAGAGATGGCAGCTGAGAAAGTATGACCCGAAAGTG 466
Db 421 GTTCTTGTCAATACCAAGCCCGAGAGATGGCAGCTGAGAAAGTATGACCCGAAAGTG 480
Oy 467 GAGCTTGAAGAGAGATGGTCTGTGTACTTCAAGAGCCCTAAGCCAGGTGGACCGAA 526
Db 481 GAGCTTGAAGAGAGATGGTCTGTGTACTTCAAGAGCCCTAAGCCAGGTGGACCGAA 540
Oy 527 ATGGCTCCAAAGCTGCATCCACATGACAAACGCAAAAGTGCCAGAGCTTGCAAGTTT 586
Db 541 ATGGCTCCAAAGCTGCATCCACATGACAAACGCAAAAGTGCCAGAGCTTGCAAGTTT 600
Oy 587 GAAGAAACAGGAATCTCTCATAGTGAATTTTCCATGCTCAGATACGGAAGAGTGT 646
Db 601 GAAGAAACAGGAATCTCTCATAGTGAATTTTCCATGCTCAGATACGGAAGAGTGT 660
Oy 647 GTGCCCTTGAGTCTCTGAAATCTCTAACCCTTGCACTTGGCTTCAAGCTGAC 706
Db 661 GTGCCCTTGAGTCTCTGAAATCTCTAACCCTTGCACTTGGCTTCAAGCTGAC 720
Oy 707 CAGGAGTTTACATGACGCTTGATTAAGAGGTGATGACATGCTTGTGCTGAGCTC 766
Db 721 CAGGAGTTTACATGACGCTTGATTAAGAGGTGATGACATGCTTGTGCTGAGCTC 780
Oy 767 TTGAGGAACTAAGAGATTTTACAGACGCTAATAATCAGAAATGTTTCCGAAATAGC 826
Db 781 TTGAGGAACTAAGAGATTTTACAGACGCTAATAATCAGAAATGTTTCCGAAATAGC 840
Oy 827 CAGGACTATCAACTGGATCTTCCAAATCAATTTGGCTTCAAGAAATTCAGAGTACTG 886
Db 841 CAGGACTATCAACTGGATCTTCCAAATCAATTTGGCTTCAAGAAATTCAGAGTACTG 900
Oy 887 ATCACTGAGGAAATGACACTGAGACTGAGTAAACAGCTTCTTAAGAAAG----- 938
Db 901 ATCACTGAGGAAATGACACTGAGACTGAGTAAACAGCTTCTTAAGAAAGATTTAG 960
Oy 939 ----- 938
Db 961 GCTCTGAACAACTAATTAAGATATGCCGGAACAAACCGATGGTTAAAAACCGT 1020
Oy 939 -----GACCTGGCCATATGTCCTCCCTGCTATGAGCTTAGAGTATGATGTC 988
Db 1021 TTTTGAACAACCTGCTGCTATGTCCTCCCTGCTATGAGCTTAGAGTATGATGTC 1080
Oy 989 TCGAAGTGGAGAGTCTGTTTGAACCTCTCTTGAATCTGCAAGTATTCATCCAG 1048
Db 1081 TCGAAGTGGAGAGTCTGTTTGAACCTCTCTTGAATCTGCAAGTATTCATCCAG 1140
Oy 1049 GGCCACAAGCCTACAGCCATCCATAAAGATGCCATCAATGAAGCTGAGAACAGAGA 1108
Db 1141 GGCCACAAGCCTACAGCCATCCATAAAGATGCCATCAATGAAGCTGAGAACAGAGA 1200
Oy 1109 AGTTATCACTGTGTGACCTGTGTGATGAATCATATGGGATCGGAAATGGCAGCG 1168
Db 1201 AGTTATCACTGTGTGACCTGTGTGATGAATCATATGGGATCGGAAATGGCAGCG 1260
Oy 1169 CACATTAATCAATCCATTTGAACCAATGAAGAAAGAAAGAAAGATTTGACTCAGAT 1228
Db 1261 CACATTAATCAATCCATTTGAACCAATGAAGAAAGAAAGAAAGATTTGACTCAGAT 1320
```

```

Db 1939 TTCAGAAATCTACAGAGAGAGGAATCAGACTGAGAGAGCTGTGACATAGGACTTG 1998
QY 1910 AAGACCAAGAGACTTTGAAATTTTGCAGAGCTCATGATGTATATATACAGCTGCT 1969
    |||||
Db 1999 AAGACCAAGAGACTTTGAAATTTTGCAGAGCTCATGATGTATATATACAGCTGCT 2058
QY 1970 TTCTATGAGTACAAATCTATATTTTATTTATTTGAGATTAAATAAGAAAAATTTTACAG 2029
    |||||
Db 2059 TTCTATGAGTACAAATCTATATTTTATTTATTTGAGATTAAATAAGAAAAATTTTACAG 2118
QY 2030 AAAAAAAAAA 2041
    |||||
Db 2119 AAAAAAAAAA 2130

RESULT 3
US-10-170-235-40554
; Sequence 40554, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 40554
; LENGTH: 2119
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-40554

Query Match          91.6%; Score 1869.6; DB 8; Length 2119;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 2008; Conservative 0; Mismatches 9; Indels 103; Gaps 3;

QY 11 ATGGCGTCCGTGGCGGCTGACAGAGAGTCTGTGGGGAGTGGGGCTCAGGGGCTGCGCAA 70
    |||||
Db 1 ATGGCGTCCGTGGCGGCTGACAGAGAGTCTGTGGGGAGTGGGGCTCAGGGGCTGCGCAA 60
QY 71 CGGACCCCTACTCTTGTAGTATTCGCGGGCCACGGGCAACCGCAATCCAGCTGGCG 130
    |||||
Db 61 CGGACCCCTACTCTTGTAGTATTCGCGGGCCACGGGCAACCGCAATCCAGCTGGCG 120
QY 131 TTGCACTAGGCGGCGGCTGCGGGTGTGAGATGTGACAGCGCTCACTTCATGCAAGTCTAT 190
    |||||
Db 121 TTGCACTAGGCGGCGGCTGCGGGTGTGAGATGTGACAGCGCTCACTTCATGCAAGTCTAT 180
QY 191 GAAGGCTAGACATCATCAACAAGATTCTGCCAAGAGAGAGAAATCTGCGGCGAC 250
    |||||
Db 181 GAAGGCTAGACATCATCAACAAGATTCTGCCAAGAGAGAGAAATCTGCGGCGAC 240
QY 251 CACATGATAGCTTTGTGTGATCTCTTGTGACCAATTACACAGTGTGAGATTGAGAAAT 310
    |||||
Db 241 CACATGATAGCTTTGTGTGATCTCTTGTGACCAATTACACAGTGTGAGATTGAGAAAT 300
QY 311 AGAGCAACTGCTCTGATTTGAA-----GATATATTGGCCGA 346
    |||||
Db 301 AGAGCAACTGCTCTGATTTGAA-----GATATATTGGCCGA 350
QY 347 GACAAATTTCTATTGTTGTGGAGAGACCAATTATTACATTGAATCTGCTCTGAGAA 406
    |||||
Db 361 GACAAATTTCTATTGTTGTGGAGAGACCAATTATTACATTGAATCTGCTCTGAGAA 420
QY 407 GTTCTTTGTAATACCAAGCCCAAGAGATGGGCACTGAGAAATGATTTGACGGAAGTG 466
    |||||
Db 421 GTTCTTTGTAATACCAAGCCCAAGAGATGGGCACTGAGAAATGATTTGACGGAAGTG 480
QY 467 GAGCTTGAAGAGAGATGCTTTGACTTCAAAACGGCTAGACCAAGGAGGAGCCAGAA 526
    |||||
Db 481 GAGCTTGAAGAGAGATGCTTTGACTTCAAAACGGCTAGACCAAGGAGGAGCCAGAA 540
QY 527 ATGGCTGCCAAGCTGCATCATGACAAACGCAAGATGGCCAGAGAGCTTGAAGTTT 586

```

```

Db 541 ATGGCTCCCAAGCTGCATCCACATGACAAACGCAAAATGGCCAGAGCTTGCAAGTTT 600
QY 587 GAAGAAACAGAAATCTCTCATAGTGAATTTCTCATGTGTACATATACGGAAGAGTGCT 646
    |||||
Db 601 GAAGAAACAGAAATCTCTCATAGTGAATTTCTCATGTGTACATATACGGAAGAGTGCT 660
QY 647 GGTCCCTTGGAGGCTCTGGAAGTCTCAACCTTGCATCCCTTGGCTTCAATGCTGAC 706
    |||||
Db 661 GGTCCCTTGGAGGCTCTGGAAGTCTCAACCTTGCATCCCTTGGCTTCAATGCTGAC 720
QY 707 CAGGCACTTCTAGATGACGCTTGGATPAGAGGATGATGACATGCTTGTGCTGGGCTC 766
    |||||
Db 721 CAGGCACTTCTAGATGACGCTTGGATPAGAGGATGATGACATGCTTGTGCTGGGCTC 780
QY 767 TTGGAGGAGTAAAGATTTTTCACAGAGCTATPACAGAGAAATGTTTCGGAATATAGC 826
    |||||
Db 781 TTGGAGGAGTAAAGATTTTTCACAGAGCTATPACAGAGAAATGTTTCGGAATATAGC 840
QY 827 CAGGCACTTCTAGATGACGCTTGGATPAGAGGATGATGACATGCTTGTGCTGGGCTC 886
    |||||
Db 841 CAGGCACTTCTAGATGACGCTTGGATPAGAGGATGATGACATGCTTGTGCTGGGCTC 900
QY 887 ATCACTAGGAGAAATACACACTGAGAGACTAGTACAGAGCTTCTTAAAGAAAG----- 938
    |||||
Db 901 ATCACTAGGAGAAATACACACTGAGAGACTAGTACAGAGCTTCTTAAAGAAAGCTATTGAG 960
QY 939 ----- 938
Db 961 GCTCTGAACAGTAACATAAGATATGCCGGAACAAACCGATGGTTTAAACCGT 1020
QY 939 -----GACCTGCTCCCATTTGTCGCCCTGTCTATGCTGTACAGATATGATGTC 988
    |||||
Db 1021 TTTTGAAGAGAGCTGCTCCATTTGCCCTGTCTATGCTGTACAGATATGATGTC 1080
QY 989 TCGAAGTGGAGAGAGTGTCTTGAACCTGCTTGAATTCGTAATTCGTAATTCATGCTCAG 1048
    |||||
Db 1081 TCGAAGTGGAGAGAGTGTCTTGAACCTGCTTGAATTCGTAATTCGTAATTCATGCTCAG 1140
QY 1049 GGCCACAAGCTACAGAGCTTCCATTAAGATGCAATACATGAGCTGAGAAACAGAGA 1108
    |||||
Db 1141 GGCCACAAGCTACAGAGCTTCCATTAAGATGCAATACATGAGCTGAGAAACAGAGA 1200
QY 1109 AGTTATCAGCTGTGTGACCTCTGTGATCAATATCATCTATGGGATGCGCAATGGCAGCG 1168
    |||||
Db 1201 AGTTATCAGCTGTGTGACCTCTGTGATCAATATCATCTATGGGATGCGCAATGGCAGCG 1260
QY 1169 CACATATAATCCAAATCCCACTTGAACCACTGAAGAAAGAAAGATGAGCTAGAT 1228
    |||||
Db 1261 CACATATAATCCAAATCCCACTTGAACCACTGAAGAAAGAAAGATGAGCTAGAT 1320
QY 1229 GCTGTCAACACATAGAAAGTCAAGATGTTTCCCAAGCTATPAAACAAGAACTAAAGG 1288
    |||||
Db 1321 GCTGTCAACACATAGAAAGTCAAGATGTTTCCCAAGCTATPAAACAAGAACTAAAGG 1380
QY 1289 AAGGATCCCCAGGCGAAGATGATCAAGAGCTGAATGCAAGCTTAAAGAGATATGCCA 1348
    |||||
Db 1381 AAGGATCCCCAGGCGAAGATGATCAAGAGCTGAATGCAAGCTTAAAGAGATATGCCA 1440
QY 1349 GTGGCTTTGGAAGAGTGTGGGAGATCCAGTTAGAGAGGAGGATATGTTTCTCCCA 1408
    |||||
Db 1441 GTGGCTTTGGAAGAGTGTGGGAGATCCAGTTAGAGAGGAGGATATGTTTCTCCCA 1500
QY 1409 GTGTGGCAAGAGATGCTATGCGGAATTCCTGCAATAGCAAGAAAGCTCCACATTTT 1468
    |||||
Db 1501 GTGTGGCAAGAGATGCTATGCGGAATTCCTGCAATAGCAAGAAAGCTCCACATTTT 1560
QY 1469 CTTTATGATGCTTTAAAGTCTCAAGTCTCTATATATGAAACACAGAGCTTGTGACG 1528
    |||||
Db 1561 CTTTATGATGCTTTAAAGTCTCAAGTCTCTATATATGAAACACAGAGCTTGTGACG 1620
QY 1529 TCCCTGTGCGATGTCTGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1588
    |||||

```


: PRIOR FILING DATE: 2000-09-11
 :
 : NUMBER OF SEQ ID NOS: 888
 :
 : SOFTWARE: Custom
 :
 : SEQ ID NO 119
 : LENGTH: 2130
 :
 : TYPE: DNA
 :
 : ORGANISM: Homo sapiens
 :
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (191)..(1426)
 :
 : US-10-380-731-119

Query Match	94.1%;	Score 1921.2;	DB 8;	Length 2130;
Best Local Similarly	95.9%;	Pred. No. 0;		
Matches 2025;	Conservative	0;	Mismatches 8;	Indels 79;
				Gaps 2

QY	8	AAGTATGGCGTCGCGGGGGGTGCAACGAGCAAGTTCCTGTGGGCGAGTGGGCTACAGGGGCTGTG	67
Db	20	ACGATTCGCGTCGCGGGGGGTGCAACGAGCAAGTTCCTGTGGGCGAGTGGGCTACAGGGGCTGTG	79
QY	68	CAAGGAGCCCTACTCTTTGTATAGATTCCTGGGGGCCAAGGGGCAACGGGCAATCCAGGCTG	127
Db	80	CAAGGAGCCCTACTCTTTGTATAGATTCCTGGGGGCCAAGGGGCAACGGGCAATCCAGGCTG	139
QY	128	GCCTTGGAGCTAGGCGCAGCGGCTGCGGGGTGAGATCGTAGCCGCTGACTCATGACGAGTC	187
Db	140	GCCTTGGAGCTAGGCGCAGCGGCTGCGGGGTGAGATCGTAGCCGCTGACTCATGACGAGTC	199
QY	188	TATGAGGCGCTAGACATCATCACCACAAAGCTTCTGCCCCAAGACGACAGAAATCTCCGG	247
Db	200	TATGAGGCGCTAGACATCATCACCACAAAGCTTCTGCCCCAAGACGACAGAAATCTCCGG	259
QY	248	CACCAATGATCAGCTTTGTGTGATCCCTGTGTGACCAATACAGAGGTGGAGCTTCAGA	307
Db	260	CACCAATGATCAGCTTTGTGTGATCCCTGTGTGACCAATACAGAGGTGGAGCTTCAGA	319
QY	308	AATAGAGCACTGCTCTGATTTGAAGATATATTTGCCGAGACAAATTCCTATTTGTTGTG	367
Db	320	AATAGAGCACTGCTCTGATTTGAAGATATATTTGCCGAGACAAATTCCTATTTGTTGTG	379
QY	368	GGAGGAAACCAATTATACATTTGATCTCTGCTCTGGAAATTTCTTGCAATACCAAGCCC	427
Db	380	GGAGGAAACCAATTATACATTTGATCTCTGCTCTGGAAATTTCTTGCAATACCAAGCCC	439
QY	428	CAGGAGTGGGCACTGTGAAGAAAGTATTGACCGAAAGTGGAGCTTGAAAGAGAGATGCT	487
Db	440	CAGGAGTGGGCACTGTGAAGAAAGTATTGACCGAAAGTGGAGCTTGAAAGAGAGATGCT	499
QY	488	CTTGACTTCAACAAACGCTTAGCCAGGTGACCCAGAAATGGCTGCCAAGCTGCATCCA	547
Db	500	CTTGACTTCAACAAACGCTTAGCCAGGTGACCCAGAAATGGCTGCCAAGCTGCATCCA	559
QY	548	CATGCAAGCAAGCAAAAGTGGCGAGGAGCTGGCAATTTTGAAGAAACAGGAATCTCTCAT	607
Db	560	CATGCAAGCAAGCAAAAGTGGCGAGGAGCTGGCAATTTTGAAGAAACAGGAATCTCTCAT	619
QY	608	AGTGAATTTCTCAATCGTCAACATACGGAAGAAGTGGTGCCTCCCTTGGAGGTCTCTG	667
Db	620	AGTGAATTTCTCAATCGTCAACATACGGAAGAAGTGGTGCCTCCCTTGGAGGTCTCTG	679
QY	668	AAGTCTCTAAACCTTCGATCCTTTGGGCTTATGCTGACCAAGCAAGTTCATGATGAGCGC	727
Db	680	AAGTCTCTAAACCTTCGATCCTTTGGGCTTATGCTGACCAAGCAAGTTCATGATGAGCGC	739
QY	728	TTGGATTAAGAGGGTGATGATCGTCTGCTGCTGAGGCTCTTGGAGGAACCTAAGAGATTTT	787
Db	740	TTGGATTAAGAGGGTGATGATCGTCTGCTGCTGAGGCTCTTGGAGGAACCTAAGAGATTTT	799
QY	788	CACACACCCCTAATACAAAGAAATGTTTCGAAAAATAGCCAGGACTATCAACATGATATC	847
Db	800	CACACACCCCTAATACAAAGAAATGTTTCGAAAAATAGCCAGGACTATCAACATGATATC	859
QY	848	TTTCAATCAATTTGGCTTCAAGGAATTTTCAGAGTATCCTGATCCTGAAGGCAAAATTCACA	907

Db	860	TTCCATTCATTGGCTTCAAGGAATTTACAGAGTACGTACGTACGTGAGGGAAATGACACA	919
Qy	908	CTGGAGACTGTACCAAGCTTCTTAAAGAAAG-----	938
Db	920	CTGGAGACTGTAAACAGCTTCTTAAAGAAAGTATTTAGAGGCTCTGAAACAACTAATAAG	979
Qy	939	-----GACCTGGGATCC	949
Db	980	AGATATGCCCCGGAACAAACCGATGGGTTTAAAAACCGTTTTTTTGAGCAGACCTGGTGCC	1039
Qy	950	ATTGTCCCCCTGTCTATGGCTTAGAGATATCGATGTCTGAAAGTGGAGAGACTGTGT	1009
Db	1040	ATTGTCCCCCTGTCTATGGCTTAGAGATATCGATGTCTGAAAGTGGAGAGACTGTGT	1099
Qy	1010	CTTGAACCTGTCTTGAATCGTGCAAGTTTCTATCCAGGGCCACAGCCCTACAGCCACT	1069
Db	1100	CTTGAACCTGTCTTGAATCGTGCAAGTTTCTATCCAGGGCCACAGCCCTACAGCCACT	1159
Qy	1070	CCATTAAGATGGCATATCATGAGAGTGGAAACAAGAGAGTTATCACCTGTGAGACTC	1129
Db	1160	CCATTAAGATGGCATATCATGAGAGTGGAAACAAGAGAGTTATCACCTGTGAGACTC	1219
Qy	1130	TGTATGCAATCATCATTTGGGGATTCGCGAATGGGCGACACATTAATCCAAATCCAC	1189
Db	1220	TGTATGCAATCATCATTTGGGGATTCGCGAATGGGCGACACATTAATCCAAATCCAC	1279
Qy	1190	TTGAACCACTGAAGAAAGAGAAGATTGAGACTCAGATGCTGTCAACACATAGAAAGT	1249
Db	1280	TTGAACCACTGAAGAAAGAGAAGATTGAGACTCAGATGCTGTCAACACATAGAAAGT	1339
Qy	1250	CAGAGTGTTCCTCCAGACTATTAACAAAGAACCTAAAGGGAAGGATCCCGACGGCAGAT	1309
Db	1340	CAGAGTGTTCCTCCAGACTATTAACAAAGAACCTAAAGGGAAGGATCCCGACGGCAGAT	1399
Qy	1310	GATCAAGAGCTGAATACACACTGTTTAAAGACATGTCCAGTGGCCTTTGAAAGGTGGTG	1369
Db	1400	GATCAAGAGCTGAATACACACTGTTTAAAGACATGTCCAGTGGCCTTTGAAAGGTGGTG	1459
Qy	1370	GGGATCCAGTTCAGAGGAGGGGTATGTTGTCTCCAGTGTGGCAAGAGAGTGTCTAT	1429
Db	1460	GGGATCCAGTTCAGAGGAGGGGTATGTTGTCTCCAGTGTGGCAAGAGAGTGTCTAT	1519
Qy	1430	GCGGAATCTCTGATATGCGAAGAAAGTCCACACTTTCTTTGATGTGGTTTAAAGT	1489
Db	1520	GCGGAATCTCTGATATGCGAAGAAAGTCCACACTTTCTTTGATGTGGTTTAAAGT	1579
Qy	1490	CTCACGTCTCTATTAATAGAAACAGCAGTCTGTGAGTCCCTGTGTGGCTGATGTGTC	1549
Db	1580	CTCACGTCTCTATTAATAGAAACAGCAGTCTGTGAGTCCCTGTGTGGCTGATGTGTC	1639
Qy	1550	TGGAATGATGTAGTTCAGGAAGCATTTTTTTTTTTCTTTGAACCTTAAAGTCTATTA	1609
Db	1640	TGGAATGATGTAGTTCAGGAAGCATTTTTTTTTTTCTTTGAACCTTAAAGTCTATTA	1699
Qy	1610	TTAAAGCAGACAGATTTCCACATTTTATACATGAGGATTCCTTTGGGTGATATACA	1669
Db	1700	TTAAAGCAGACAGATTTCCACATTTTATACATGAGGATTCCTTTGGGTGATATACA	1759
Qy	1670	GGATTGACGTGATCCCTTTTAAAGAAATTTTATGTGCCCTGACTGTGGCTTAAATATCTA	1729
Db	1760	GGATTGACGTGATCCCTTTTAAAGAAATTTTATGTGCCCTGACTGTGGCTTAAATATCTA	1818
Qy	1730	ATTTCACAGATCTTTTGTAGATGACTGAATATTTGTGAGCCACATATTGGGATGTTCTAG	1789
Db	1819	ATTTCACAGATCTTTTGTAGATGACTGAATATTTGTGAGCCACATATTGGGATGTTCTAG	1878
Qy	1790	ATTGATGATGATGCGAGAAAGGCCATCTCCACTTGAGATGATTAAAGTAAACCAACTAG	1849
Db	1879	ATTGATGATGATGCGAGAAAGGCCATCTCCACTTGAGATGATTAAAGTAAACCAACTAG	1938
Qy	1850	TTTCGGAATCTACAGAAAGGGGAATCAGACTGAGGAAGCTGTGACATAGGACTTG	1909

140 GCGTTCAGCTAGCGAGCGGCTCGCGGTGAGATGCTGACGCTGATCCATGACAGTCC 199
188 TAGAAGGCGCTAGACATCATCACCACCAAGCTTCTGCCCAAGACGAGAAATCTGGCGG 247
200 TAGAAGGCGCTAGACATCATCACCACCAAGCTTCTGCCCAAGACGAGAAATCTGGCGG 259
248 CACACATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
260 CACACATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 319
308 AATAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367
320 AATAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379
368 GAGAGAACCAATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 427
380 GAGAGAACCAATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 439
428 CAGAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 487
440 CAGAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 499
488 CTGCTACTTCAAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547
500 CTGCTACTTCAAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 559
548 CAGAGAACCAATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 607
560 CAGAGAACCAATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 619
608 AGTGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 667
620 AGTGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 679
668 AAGTCTCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 727
680 AAGTCTCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 739
728 TTGATTAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787
740 TTGATTAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 799
788 CACAGAGCTATATATCAGAAAGATGTTTCGAAAAATAGCAGACTATCAACATGATGATC 847
800 CACAGAGCTATATATCAGAAAGATGTTTCGAAAAATAGCAGACTATCAACATGATGATC 859
848 TTCCATCAATTTGGCTTCAAGAAATTTACGAGTACTGATCAGAGGCAAAATGACACA 907
860 TTCCATCAATTTGGCTTCAAGAAATTTACGAGTACTGATCAGAGGCAAAATGACACA 919
908 CTGAGAGCTAGTAACCGCTTCTAAGAAG----- 938
920 CTGAGAGCTAGTAACCGCTTCTAAGAAGAGTATGAGGCTCTGAAAACAAGTAAG 979
939 -----GACTGCTGCC 949
980 AGATATCCCGGAAACAAACCGATGGGTTAAAAACCGTTTTCGACAGACCTGGCTGCC 1039
950 ATTGTCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1009
1040 ATTGTCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1099
1010 CTGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1069
1100 CTGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1159
1070 CCAATAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1129
1160 CCAATAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1219
1130 TGTGATGAATCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1189

1220 TGTGATGAATCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1279
1190 TTGAACCACTGAGAAAGAAAGAGATGAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1249
1280 TTGAACCACTGAGAAAGAAAGAGATGAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1339
1250 CAGAGTCTTCCCGACACTATTAACAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1309
1340 CAGAGTCTTCCCGACACTATTAACAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1399
1310 GATCAGAGCTGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1369
1400 GATCAGAGCTGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1459
1370 GGGATCCAGTTCAG 1429
1460 GGGATCCAGTTCAG 1519
1430 GCGGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1489
1520 GCGGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1579
1490 CTCACGTTCTCTATTAATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1549
1580 CTCACGTTCTCTATTAATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1639
1550 TGGAAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1609
1640 TGGAAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1659
1610 TTAAGACAGACAGATTCACATTTTATACATGAGAGATCTTCTTCTGGAACCTTAAAGTTCT 1669
1700 TTAAGACAGACAGATTCACATTTTATACATGAGAGATCTTCTTCTGGAACCTTAAAGTTCT 1759
1670 GGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1729
1760 GGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1818
1730 ATTCCAGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1789
1819 ATTCCAGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1878
1790 ATTGAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1849
1879 ATTGAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1938
1850 TTCTCGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1909
1939 TTCTCGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1998
1910 AAGACCAAGACTTGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1969
1999 AAGACCAAGACTTGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2058
1970 TTCTATGAGTTACAAATCTATATTTTATGAGATTTAAATAAAGAAAAATTTACAG 2029
2059 TTCTATGAGTTACAAATCTATATTTTATGAGATTTAAATAAAGAAAAATTTACAG 2118
2030 AAAAAAAAAA 2041
2119 AAAAAAAAAA 2130

RESULT 2
US-10-380-731-119
; Sequence 119, Application US/10380731
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-114
; CURRENT APPLICATION NUMBER: US/10/380, 731
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 09/659, 671

QY 1219 GGACTCAGATGCTGTCAACACCATAGAAAAGTCAGAGTGTTCCTCCAGACTATATACAAAGA 1278
|||||
Db 413 GGACTCAGATGCTGTCAACACCATAGAAAAGTCAGAGTGTTCCTCCAGACTATATACAAAGA 472
QY 1279 ACCATAAGGGAAGGATCCCGAGGAGAAATGATCAAGAGTGAATGCAGCGTTTAAAGA 1338
|||||
Db 473 ACCATAAGGGAAGGATCCCGAGGAGAAATGATCAAGAGTGAATGCAGCGTTTAAAGA 532
QY 1339 GACATGTCCAGTGGCTTTGGAAGGTGTGGGATCCAGTTCAGAGGAGGGGTATGT 1398
|||||
Db 533 GACATGTCCAGTGGCTTTGGAAGGTGTGGGATCCAGTTCAGAGGAGGGGTATGT 592
QY 1399 TTGTCTCCAGTGGGCAAGAGAGTGTATGGGAAATCTCGCAATAGCAGAAAAGCTC 1458
|||||
Db 593 TTGTCTCCAGTGGGCAAGAGAGTGTATGGGAAATCTCGCAATAGCAGAAAAGCTC 652
QY 1459 CCACATTTTCTTTTATGATGTGTTTAAAGTCTCAGCTTCTATATAGAAAACAGCAG 1518
|||||
Db 653 CCACATTTTCTTTTATGATGTGTTTAAAGTCTCAGCTTCTATATAGAAAACAGCAG 712
QY 1519 TCTTGTACGCTCCTTGTGTGTGTATGTCTGGAATGATGTAGTTACAGAAAAGCA -TT 1577
|||||
Db 713 TCTTGTACGCTCCTTGTGTGTGTATGTCTGGAATGATGTAGTTACAGAAAAGCA -TT 772
QY 1578 TTTTTCCTTTGAACCTTAAAGTCTATATATTAAGCAGACAGATCCACATTTT 1637
|||||
Db 773 TTTTTCCTTTGAACCTTAAAGTCTATATATTAAGCAGACAGATCCACATTTT 832
QY 1638 ATACATGAGAGATCTCTTTGTGTGAATACACAGATGATGCATCCCTTTAAAGAGT 1697
|||||
Db 833 ATACATGAGAGATCTCTTTGTGTGAATACACAGATGATGCATCCCTTT -AAAGAGT 891
QY 1698 TTTATGTCCTGACCTGCTGCTAAATATCTAAATTCAGAGATGCTTTGTAGATGCTGA 1757
|||||
Db 892 TTTATGTCCTGACCTGCTGCTAAATATCTAAATTCAGAGATGCTTTGTAGATGCTGA 951
QY 1758 AGTATTTGTGAGCCACATATTTGGAGTCTAGATTTGAGTAATGCAGAGAAAGGCCAT 1817
|||||
Db 952 AGTATTTGTGAGCCACATATTTGGAGTCTAGATTTGAGTAATGCAGAGAAAGGCCAT 1011
QY 1818 CTCATTTGAGATGATTAAGTGAACCAACTAGTTCGGAATTCACAGAGAAAGAGGA 1877
|||||
Db 1012 CTCATTTGAGATGATTAAGTGAACCAACTAGTTCGGAATTCACAGAGAAAGAGGA 1071
QY 1878 ATCAGACTGAGGAAGCTGTGACATAGGACTTGAAGACCAAGACTTTGAAATTTGCGAGC 1937
|||||
Db 1072 ATCAGACTGAGGAAGCTGTGACATAGGACTTGAAGACCAAGACTTTGAAATTTGCGAGC 1131
QY 1938 TGCTCAATGCTGATTAATTAATCACTGCTTCTATGAGTTACAAATCTATATTTT 1997
|||||
Db 1132 TGCTCAATGCTGATTAATTAATCACTGCTTCTATGAGTTACAAATCTATATTTT 1191
QY 1998 ATTGAAGTTTAAATAAGAAAAAATTACAG 2029
|||||
Db 1192 ATTGAAGTTTAAATAAGAAAAAATTACAG 1223

Search completed: April 21, 2003, 17:14:09
Job time : 6044.18 secs

QY	956	CCCCCTGTATGCGCTTAAAGATATCGATGTCTGAGTGGAGGAGCTGTCTTGA	1015
Db	661	CCCCCTGTACGCGCTTAAAGATATCGATGTCTGAGTGGAGGAGCTGTCTTGA	720
QY	1016	CCCTCTTAAATCGCGAAAGTTTCATCCAGGGGCGCAACCCACACCTCCAA	1075
Db	721	CTCTCTTTAAATCGCGAAAGTTTCATCCAGGGGCGCAACCCACACCTCCAA	780
QY	1076	AAGATGCCATACAAATGAAGCTGGAACAAGAGAAGTTATCACCTGTGACCTGTGAT	1135
Db	781	AAGATGCCATACAAATGAAGCTGGAACAAGAGAAGTTATCACCTGTGACCTGTGAT	840
QY	1136	CGATCATCATTTGGGGATGCCGAATGGGAGGGCCATATAATCCAAATCCCACTTGAC	1195
Db	841	CGAATCATCATTTGGGGATGCCGAATGGGAGGGCCATATAATCCAAATCCCACTTGAC	900
QY	1196	CAACTGAAGAAAAGAAAGAAATTTGGACATGATGTCTGCATACACCTAGAAAGTCAGT	1255
Db	901	CAACTGAAGAAAAGAAAGAAATTTGGACATGATGTCTGCATACACCTAGAAAGTCAGT	960
QY	1256	GTTTCCCCAGACTTAAACAAGACCTAAAGGAGGGATCCCAAGGCGAGATGATCA	1315
Db	961	GTTTCCCCAGACTTAAACAAGACCTAAAGGAGGGATCCCAAGGCGAGATGATCA	1020
QY	1316	GAGCTGAATATGCACGCTTTAAGAGACATGCCAGTGGCCTTTGGAAAAGTGGTGGGATC	1375
Db	1021	GAGCTGAATATGCACGCTTTAAGAGACATGCCAGTGGCCTTTGGAAAAGTGGTGGGATC	1080
QY	1376	CAGTTCAAGAGGAGGGGATGTGTGTCTCCAGCTGGGCAAGAGAGTCTATCGGA	1435
Db	1081	CAGTTCAAGAGGAGGGGATGTGTGTCTCCAGCTGGGCAAGAGAGTCTATCGGA	1140
QY	1436	TTCTCTGCATAGCAAGAAACCTCCACCATTTCTTTATGATGGTTTAAAGTCACG	1495
Db	1141	TTCTCTGCATAGCAAGAAACCTCCACCATTTCTTTATGATGGTTTAAAGTCACG	1200
QY	1496	TTCTCTATAATAGAAACAGCAGGCTGTGTGACCTGTGTGGCTGATGTCTGAAA	1555
Db	1201	TTCTCTATAATAGAAACAGCAGGCTGTGTGACCTGTGTGGCTGATGTCTGAAA	1260
QY	1556	TGATGTAGTTCAGGAAGCAATTTTTTTTCTTGAACCTTAAAGTCTATATTAAA	1615
Db	1261	TGATGTAGTTCAGGAAGCAATTTTTTTTCTTGAACCTTAAAGTCTATATTAAA	1320
QY	1616	GCACACAGATTCACACATTTTTTATACATGAGATCTTCTTGTGTGAATACAGAGATG	1675
Db	1321	GCACACAGATTCACACATTTTTTATACATGAGATCTTCTTGTGTGAATACAGAGATG	1380
QY	1676	ACTGCATCCCTTTAAAGAAGTTTATGTCCCTGCACTGTGGCTAAATATCTAATTCC	1735
Db	1381	ACTGCATCCCTTTAAAGAAGTTTATGTCCCTGCACTGTGGCTAAATATCTAATTCC	1439
QY	1736	AGATGCTTTTATAGATACATGAAGTATTTTGTAGCCACATATTGGAGTGTAGATTGA	1795
Db	1440	AGATGCTTTTATAGATACATGAAGTATTTTGTAGCCACATATTGGAGTGTAGATTGA	1499
QY	1796	GTTGAATGGCAAGAAAGGCCATCTCCATTGAGATGATTAAGTGAACCAACTAGTCTCG	1855
Db	1500	GTTGAATGGCAAGAAAGGCCATCTCCATTGAGATGATTAAGTGAACCAACTAGTCTCG	1559
QY	1856	GAAATTCACAGAAAGAGGAGATTCAGATCGAGAGAGCTGACATAGGACTTGAAGCC	1915
Db	1560	GAAATTCACAGAAAGAGGAGATTCAGATCGAGAGAGCTGACATAGGACTTGAAGCC	1619
QY	1916	AAAGACTTTGAATTTCCGAGCTGCATGTGGAATTTATATCACTGCTGCTTCTAT	1975
Db	1620	AAAGACTTTGAATTTCCGAGCTGCATGTGGAATTTATATCACTGCTGCTTCTAT	1679
QY	1976	TGAGTTACAAATCTATATTTTATATGAAGTTTAAATTAAGAAAATTTTACAGAAAAA	2035
Db	1680	TGAGTTACAAATCTATATTTTATATGAAGTTTAAATTAAGAAAATTTTACAGAAAAA	1739

QY	2036	AAAAAA	2041
Db	1740	AAAAAA	1745
<p>RESULT 12</p> <p>US-09-757-028-874</p> <p>Sequence 874, Application US/09757028</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Rosen et al.</p> <p>TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies</p> <p>FILE REFERENCE: PM001</p> <p>CURRENT FILING DATE: 2001-01-09</p> <p>PRIOR APPLICATION NUMBER: 60/179,065</p> <p>PRIOR FILING DATE: 2000-01-31</p> <p>PRIOR APPLICATION NUMBER: 60/180,628</p> <p>PRIOR FILING DATE: 2000-02-04</p> <p>NUMBER OF SEQ ID NOS: 2660</p> <p>SOFTWARE: PatentIn Ver. 2.0</p> <p>SEQ ID NO 874</p> <p>LENGTH: 1183</p> <p>TYPE: DNA</p> <p>ORGANISM: Homo sapiens</p> <p>FEATURE:</p> <p>NAME/KEY: SITE</p> <p>LOCATION: (1172)</p> <p>OTHER INFORMATION: n equals a,t,g, or c</p> <p>NAME/KEY: SITE</p> <p>LOCATION: (1177)</p> <p>OTHER INFORMATION: n equals a,t,g, or c</p> <p>US-09-757-028-874</p>			
QY	919	TACACAGCTTTAAAGAAAGACACCTGGTCCCATTTGGCCCCCGTGTATGAGGTAGGT	978
Db	15	TAAAAACGCTTTTGTGAGCAGACCTGGTCCCATTTGGCCCCCGTGTATGAGGTAGGT	74
QY	979	ATCTGATGCTCGAAGTGGAGAGTCTGTCTTGAACCTGCTTTGAATCGTGC	1038
Db	75	ATCTGATGCTCGAAGTGGAGAGTCTGTCTTGAACCTGCTTTGAATCGTGC	134
QY	1039	TTTCATCCAGGGCCACACACCTACGACCCACTCCCAATGAAGATGCCATTAATGAAGCTCA	1098
Db	135	TTTCATCCAGGGCCACACACCTACGACCCACTCCCAATGAAGATGCCATTAATGAAGCTCA	194
QY	1099	GAAACAAGAAAGTTATCACCCTGTGTACCTGTGTGATCGAATCATCATTTGGGGATCGCA	1158
Db	195	GAAACAAGAAAGTTATCACCCTGTGTACCTGTGTGATCGAATCATCATTTGGGGATCGCA	254
QY	1159	ATGGGAGCGCACATTAATTCCAATCCCACTTGAACCACTGAAGAAAAAGAAAGATT	1218
Db	255	ATGGGAGCGCACATTAATTCCAATCCCACTTGAACCACTGAAGAAAAAGAAAGATT	314
QY	1219	GGAGTCAGATGCTGTCAACACCATAGAAAGTCAAGAGTCTTCCCAAGATATATACAAAGA	1278
Db	315	GGAGTCAGATGCTGTCAACACCATAGAAAGTCAAGAGTCTTCCCAAGATATATACAAAGA	374
QY	1279	ACCTAAGGAAGGATCCCGAGGGGAGATGATCAAGAGTGAATGACAGGCTTAAGA	1338
Db	375	ACCTAAGGAAGGATCCCGAGGGGAGATGATCAAGAGTGAATGACAGGCTTAAGA	434
QY	1339	GACATGTCAGTGGCTTTGGAAAGTGTGGGATCCAGTTCCAGAGGAGGATGAT	1398
Db	435	GACATGTCAGTGGCTTTGGAAAGTGTGGGATCCAGTTCCAGAGGAGGATGAT	494
QY	1399	TTGTCTCCCAAGTCTGGGCAAGAGAGTGTATCGCGAATTTCTGCATAGCAAAAGCTC	1458
Db	495	TTGTCTCCCAAGTCTGGGCAAGAGAGTGTATCGCGAATTTCTGCATAGCAAAAGCTC	554

Db 721 CCTGCTTGAATGCTGCAAAAGTTTCATCCAGGCGCACAGCCCTACACCCACTCCATTA 780
 QY 1076 AAGATGCATACAAAGAGCTGAGAACAGAGATATACCTGTGTGTGTGTGTGTGTGTGTGT 1135
 Db 781 AAGATGCATACAAAGAGCTGAGAACAGAGATATACCTGTGTGTGTGTGTGTGTGTGTGT 840
 QY 1136 CGAATCATCATTTGGGGATGCGCAATGGGCGAGCGCACATTAATAATCCCACTTGTAAC 1195
 Db 841 CGAATCATCATTTGGGGATGCGCAATGGGCGAGCGCACATTAATAATCCCACTTGTAAC 900
 QY 1196 CAATGAGAGAGAGAGAGAGATTTGAGCTGCTGCTGCAACACCATAGAGAGAGAGAGAG 1255
 Db 901 CAATGAGAGAGAGAGAGAGATTTGAGCTGCTGCTGCAACACCATAGAGAGAGAGAGAG 960
 QY 1256 GTTTCGCCAGACTATTAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1315
 Db 961 GTTTCGCCAGACTATTAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 QY 1316 GAGCTGAATGAG 1375
 Db 1021 GAGCTGAATGAG 1080
 QY 1376 CAGTTGAG 1435
 Db 1081 CAGTTGAG 1140
 QY 1436 TTCTCTGCTAGCAG 1495
 Db 1141 TTCTCTGCTAGCAG 1200
 QY 1496 TTCTCTGCTAGCAG 1555
 Db 1201 TTCTCTGCTAGCAG 1260
 QY 1556 TGATGAGTTCAG 1615
 Db 1261 TGATGAGTTCAG 1320
 QY 1616 GCAGCAGAGATTCACATTTTATACATAGAGATCTTCTTGTGTGTGTGTGTGTGTGTGT 1675
 Db 1321 GCAGCAGAGATTCACATTTTATACATAGAGATCTTCTTGTGTGTGTGTGTGTGTGTGT 1380
 QY 1676 ACTGCATCCCTTAAAG 1735
 Db 1381 ACTGCATCCCTTAAAG 1439
 QY 1736 AGATGCTTTTGTAGATGAGTGAATTTGTGAGCCACATATTTGGAGTTCATGATTTGA 1795
 Db 1440 AGATGCTTTTGTAGATGAGTGAATTTGTGAGCCACATATTTGGAGTTCATGATTTGA 1499
 QY 1796 GTGATGCGAG 1855
 Db 1500 GTGATGCGAG 1559
 QY 1856 GAATTCCTAG 1915
 Db 1560 GAATTCCTAG 1619
 QY 1916 AAGACTTTGAATTTGAGAGTGCATGATGAGTGAATTTGATGATGATGATGATGATGAT 1975
 Db 1620 AAGACTTTGAATTTGAGAGTGCATGATGAGTGAATTTGATGATGATGATGATGATGAT 1679
 QY 1976 TGAGTTACAAATCTATATTTTATGAGAGTTTAAATTAAGAGAGAGAGAGAGAGAGAG 2035
 Db 1680 TGAGTTACAAATCTATATTTTATGAGAGTTTAAATTAAGAGAGAGAGAGAGAGAGAG 1739
 QY 2036 AAAAAA 2041
 Db 1740 AAAAAA 1745

; Sequence 1692, Application US/10172118
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yudong
 ; APPLICANT: Linsley, Peter
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Roberts, Chris
 ; APPLICANT: Van 't Veer, Laura
 ; APPLICANT: Van de Vijver, Marc
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-175-999
 ; CURRENT APPLICATION NUMBER: US/10/172.118
 ; CURRENT FILING DATE: 2002-06-14
 ; PRIOR APPLICATION NUMBER: 60/380,770
 ; PRIOR FILING DATE: 2002-05-14
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 1692
 ; LENGTH: 1749
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: NM_017646
 ; DATABASE ENTRY DATE: 2001-06-18
 ; US-10-172-118-1692
 Query Match 76.3%; Score 1556.8; DB 41; Length 1749;
 Best Local Similarity 95.1%; Pred. No. 0;
 Matches 1660; Conservative 0; Mismatches 7; Indels 79; Gaps 2;
 QY 374 ACCAATTATTAATTGAATTCCTGCTGGAAGTCTTGTCAATACCAAGCCGAGAG 433
 Db 1 ACCAATTATTAATTGAATTCCTGCTGGAAGTCTTGTCAATACCAAGCCGAGAG 60
 QY 434 ATGGGCACTGAGAAAGTGAATGACCGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 493
 Db 61 ATGGGCACTGAGAAAGTGAATGACCGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 120
 QY 494 CTTCAACAAAGCCTTAAGCCAGAGTGGACCCAGAAATGGCTGCCAAGTGCATCCATGAC 553
 Db 121 CTTCAACAAAGCCTTAAGCCAGAGTGGACCCAGAAATGGCTGCCAAGTGCATCCATGAC 180
 QY 554 AAACGGAAGTGGCCAG 613
 Db 181 AAACGGAAGTGGCCAG 240
 QY 614 TTCTTCATCGTCAACATACGGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 673
 Db 241 TTCTTCATCGTCAACATACGGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
 QY 674 TCTAACCTTGGATCTTGGCTTCACTGCTGACACGAGCAGTTCTAGATGAGCGCTTGAT 733
 Db 301 TCTAACCTTGGATCTTGGCTTCACTGCTGACACGAGCAGTTCTAGATGAGCGCTTGAT 360
 QY 734 AAGAGGTGATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
 Db 361 AAGAGGTGATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 794 CGCTATATACAGAGATTTTTCGAAATATACCAAGAGTATCAACATGATGATGATGATGATGAT 853
 Db 421 CGCTATATACAGAGATTTTTCGAAATATACCAAGAGTATCAACATGATGATGATGATGATGAT 480
 QY 854 TCAATTTGGCTCAAGAGATTTTTCGAGAGTACGTGATCAGAGAGAGAGAGAGAGAGAGAGAG 913
 Db 481 TCAATTTGGCTCAAGAGATTTTTCGAGAGTACGTGATCAGAGAGAGAGAGAGAGAGAGAGAG 540
 QY 914 ACTAGTAACCAAGCTTCTTAAGAGAG----- 939
 Db 541 ACTAGTAACCAAGCTTCTTAAGAGAG----- 600
 QY 940 -----ACCTGCTCCCATTTGTC 955
 Db 601 GCCCGGAACAAACCGATGGGTTAAAAACCGTTTTTTGAGCAGACCTGCTCCATTTGTC 660


```

Db 1571 TAATAGAAACACAGAGTCTTGTACAGCTCCTTGTGTGCTGATGCTGTGAAATGATGTA 1630
QY 1563 GTTCAGGAAGAGATTTTTTTTCTTTGAACTTAAGSTTCTATATTAAGACGAC 1622
Db 1631 GTTCAGGAAGAGATTTTTTTTCTTTGAACTTAAGSTTCTATATTAAGACGAC 1690
QY 1623 AGATTCACATTTTATACATGAGATCTTCTTTGTGTGAATACAGAGATTGACAT 1682
Db 1691 AATTCACATTTTATACATGAGATCTTCTTTGTGTGAATACAGAGATTGACAT 1750
QY 1683 CCGTTTAAAGAGTTTATGTCCCTGACCTGTGCTAAATATCTAATTCAGATGCT 1742
Db 1751 CCGTTT-AAAGAGTTTATGTCCCTGACCTGTGCTAAATATCTAATTCAGATGCT 1809
QY 1743 TTTGTAGATGATGAGATTTTGTGAGCACAATTTGGAGTCTGTGATTTGATGAAG 1802
Db 1810 TTTGTAGATGATGAGATTTTGTGAGCACAATTTGGAGTCTGTGATTTGATGAAG 1869
QY 1803 GGAGAAAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAACTAGTCTCGAATCT 1862
Db 1870 GGAGAAAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAACTAGTCTCGAATCT 1929
QY 1863 ACAGAGAGAGAGGAATCAGACTGAGAAAGCTGTGACATAGACATTTGAAGCAAGACT 1922
Db 1930 ACAGAGAGAGAGGAATCAGACTGAGAAAGCTGTGACATAGACATTTGAAGCAAGACT 1989
QY 1923 TTGAATTTGCGAGCTGCTCATGTGTGATTTATCAGCTGCTTCTATTTGATGA 1982
Db 1990 TTGAATTTGCGAGCTGCTCATGTGTGATTTATCAGCTGCTTCTATTTGATGA 2049
QY 1983 CAAATCTATTTTATTTATGAAGTTTAAATGAAGAAATTTTCAAGAAA 2034
Db 2050 CAAATCTATTTTATTTATGAGTTTAAATGAAGAAATTTTCAAGAAA 2101

```

RESULT 9 US-09-471-275-2814

Sequence 2814, Application US/09471275

GENERAL INFORMATION:

APPLICANT: Hysq, Inc.

TITLE OF INVENTION: Novel Contigs Obtained

TITLE OF INVENTION: From Various Libraries

FILE REFERENCE: 782

CURRENT APPLICATION NUMBER: US/09/471, 275

CURRENT FILING DATE: 1999-12-23

EARLIER APPLICATION NUMBER: US 09/235, 076

EARLIER FILING DATE: 1999-01-20

EARLIER APPLICATION NUMBER: US 09/234, 611

EARLIER FILING DATE: 1999-01-22

EARLIER APPLICATION NUMBER: US 09/240, 371

EARLIER FILING DATE: 1999-01-29

EARLIER APPLICATION NUMBER: US 09/277, 227

EARLIER FILING DATE: 1999-03-25

EARLIER APPLICATION NUMBER: US 09/271, 490

EARLIER FILING DATE: 1999-03-18

EARLIER APPLICATION NUMBER: US 09/293, 972

EARLIER FILING DATE: 1999-04-15

EARLIER APPLICATION NUMBER: US 09/274, 861

EARLIER FILING DATE: 1999-03-23

EARLIER APPLICATION NUMBER: US 60/125, 453

EARLIER FILING DATE: 1999-03-19

EARLIER APPLICATION NUMBER: US 60/126, 605

EARLIER FILING DATE: 1999-03-26

EARLIER APPLICATION NUMBER: US 09/306, 350

EARLIER FILING DATE: 1999-05-07

EARLIER APPLICATION NUMBER: US 09/399, 720

EARLIER FILING DATE: 1999-09-21

EARLIER APPLICATION NUMBER: US 09/404, 284

EARLIER FILING DATE: 1999-09-21

EARLIER APPLICATION NUMBER:

EARLIER FILING DATE: 1999-12-16

NUMBER OF SEQ ID NOS: 10451

SOFTWARE: pt_CT_genes Version 1.0

```

; SEQ ID NO 2814
; LENGTH: 1858
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1843)...(1283)
; OTHER INFORMATION: similar to g11419759 in the genpept database release 114,
; OTHER INFORMATION: Run with FASTX 3.3t00, default parameters
US-09-471-275-2814

```

Query Match 80.6%; Score 1646; DB 18; Length 1858;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 1773; Conservative 0; Mismatches 5; Indels 84; Gaps 4;

```

QY 12 TGGCGTCCGTCGGCGCTGACAGAGAGTCTCTGTGGGAGTGCGCTCAGGGGCTGCAAC 71
Db 3 TCGTGTCCGTCGGCGCTGACAGAGAGTCTCTGTGGGAGTGCGCTCAGGGGCTGCAAC 62
QY 72 GGACCTACCTCTTGTAGTATTCGCGGGCCAGGGGCAACGCAATCCAGCTGGGAT 131
Db 63 GCGCCCTACCTCTTGTAGTATTCGCGGGCCAGGGGCAACGCAATCCAGCTGGGAT 122
QY 132 TCGAGCTAGCGCAGCGGCTCGCGGCTGAGATGTCAGCGCTGACTCATGACAGTCTATG 191
Db 123 TCGAGCTAGCGCAGCGGCTCGCGGCTGAGATGTCAGCGCTGACTCATGACAGTCTATG 182
QY 192 AAGGCTACATCATCATCACCAACAGGTTCTTGCCCAAGAGAGAGATCTGCGGCACC 251
Db 183 AAGGCTACATCATCATCACCAACAGGTTCTTGCCCAAGAGAGAGATCTGCGGCACC 242
QY 252 ACATGATCAGCTTGTGATTCCTGTGACCAATTCACAGTGGTGGACCTCAGAAATA 311
Db 243 ACATGATCAGCTTGTGATTCCTGTGACCAATTCACAGTGGTGGACCTCAGAAATA 302
QY 312 GAGCAACTGCTGATTTAGATATATTTCCCGAGACAAATTCCTATTTGTGGAG 371
Db 303 GAGCAACTGCTGATTTAGATATATTTCCCGAGACAAATTCCTATTTGTGGAG 362
QY 372 GAACCATTTATACATTTGATGATCTCTGCTGGAAGTTCTTGTCAATACCAAGCCGACG 431
Db 363 GAACCATTTATACATTTGATGATCTCTGCTGGAAGTTCTTGTCAATACCAAGCCGACG 422
QY 432 AGATGGGCACTAGAAAGATTTGACCCGAAAGTGGAGGTTGAAAAGAGAGTGGGCTTG 491
Db 423 AGATGGGCACTAGAAAGATTTGACCCGAAAGTGGAGGTTGAAAAGAGAGTGGGCTTG 482
QY 492 TACTTACAAAGCCTTAACCCAGAGTGGAGCCAGAAATGCGTCCAAAGCTGATCCACATG 551
Db 483 TACTTACAAAGCCTTAACCCAGAGTGGAGCCAGAAATGCGTCCAAAGCTGATCCACATG 542
QY 552 ACNAAAGCAAGTGGCCAGAGCTTGCAGATTTTGTGAAGAAACAGGAATCTCTATAGTG 611
Db 543 ACNAAAGCAAGTGGCCAGAGCTTGCAGATTTTGTGAAGAAACAGGAATCTCTATAGTG 602
QY 612 AATTTTCATGCTCAACATACGGAAGAGTGGTCCCTTGGAGTCTCTGAGT 671
Db 603 AATTTTCATGCTCAACATACGGAAGAGTGGTCCCTTGGAGTCTCTGAGT 662
QY 672 TCTCTAACCTTGATCCCTTGGCTTCATGCTGACAGCAGCAGTCTAGTAGAGGCTTGG 731
Db 663 TCTCTAACCTTGATCCCTTGGCTTCATGCTGAC---CACTTCTAGTAGAGGCTTGG 718
QY 732 ATTAGAGGTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791
Db 719 ATTAGAGGTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 778
QY 792 GACGCTATATCAGAAAGATTTTGGAAATACGAGAGCTATCAACATGATCTTCC 851
Db 779 GACGCTATATCAGAAAGATTTTGGAAATACGAGAGCTATCAACATGATCTTCC 838
QY 852 AATCATTTGGCTTCAAGGAATTTCAAGATGACTGATCTGATGAGGAAATACACAGTGG 911

```


QY 1967 TCTTCTATGAGTTACAAATCTATATTTTATGAAGTTAAATAGAAAAATTAC 2026
 |||||
 Db 2042 TCTTCTATGAGTTACAAATCTATATTTTATGAAGTTAAATAGAAAAATTAC 2101
 |||||
 QY 2027 AAGAAAA 2034
 |||||
 Db 2102 AAGAAAA 2109

RESULT 8

US-60-172-360-26034
 ; Sequence 26034, Application US/60172360
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, MacDonald
 ; APPLICANT: Lal, Precl
 ; TITLE OF INVENTION: Method for the identification of Sequence Polymorphisms Using
 ; FILE REFERENCE: POLYNucleotide Sequence Databases, and Single Nucleotide Polymorph
 ; CURRENT APPLICATION NUMBER: US/60/172,360
 ; NUMBER OF SEQ ID NOS: 29838
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 26034
 ; LENGTH: 2101
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 330917.13
 ; NAME/KEY: unsure
 ; LOCATION: 2098
 ; OTHER INFORMATION: a, t, c, g, or other
 US-60-172-360-26034

Query Match 92.8% Score 1894.6; DB 61; Length 2101;
 Best Local Similarity 95.9%; Pred. No. 0;
 Matches 2007; Conservative 0; Mismatches 5; Indels 80; Gaps 3;

QY 22 GGGCGCTGCAGCAGCAGTCTCTGCGGCGAGGCGCTCAGGCGCTGCAAGGACCTACC 81
 |||||
 Db 11 GGGCGCTGCAGCAGCAGTCTCTGCGGCGAGGCGCTCAGGCGCTGCAAGGACCTACC 70
 |||||
 QY 82 TCTTGTAGTATCTCCGCGGCGAGGCGAGCGGCAATCCACGCTGCGTTCAGTAGG 141
 |||||
 Db 71 TCTTGTAGTATCTCCGCGGCGAGGCGAGCGGCAATCCACGCTGCGTTCAGTAGG 130
 |||||
 QY 142 CCAGCGCTGCGGCGAGTATGTCAGCGCTGACTCCATGCAAGGTATGAGAGCTAGA 201
 |||||
 Db 131 CCAGCGCTGCGGCGAGTATGTCAGCGCTGACTCCATGCAAGGTATGAGAGCTAGA 190
 |||||
 QY 202 CATCATGACCAACAAGGTTCTGCCAAGAGAGAGATCGCCGAGCCACATGATCAG 261
 |||||
 Db 191 CATCATGACCAACAAGGTTCTGCCAAGAGAGATCGCCGAGCCACATGATCAG 250
 |||||
 QY 262 CTTTGTGATCTCTGTCGACCAATTCACAGTGTGAGTTCAGAAATAGAGCACTGC 321
 |||||
 Db 251 CTTTGTGATCTCTGTCGACCAATTCACAGTGTGAGTTCAGAAATAGAGCACTGC 310
 |||||
 QY 322 TCTTATGAGATATATTTGGCCGAGACAAAATTCCTATTGTTGGAGAGACCAATTA 381
 |||||
 Db 311 TCTTATGAGATATATTTGGCCGAGACAAAATTCCTATTGTTGGAGAGACCAATTA 370
 |||||
 QY 382 TTACATGAAATCTGCTGTCGGAAGTCTGTCAATACCAAGCCCA-GAGATGGGCA 440
 |||||
 Db 371 TTACATGAAATCTGCTGTCGGAAGTCTGTCAATACCAAGCCCAAGGAGATGGGCA 430
 |||||
 QY 441 CTGAGAAAGTATGACCGAAGAGTGTGAAAGAGAGATGCTTACTTACA 500
 |||||
 Db 431 CTGAGAAAGTATGACCGAAGAGTGTGAAAGAGAGATGCTTACTTACA 490
 |||||
 QY 501 AACGCTTAAGCAGAGTGAGCCAGAAATGCTGCCAAGCTGCATCATGACCAACGCA 560

Db 491 AAGGCTTAAGCAGAGTGAGCCAGAAATGCTGCCAAGCTGCATCCATGACCAACGCA 550
 |||||
 QY 561 AAGTGCCAGAGAGCTTCAAGTCTTGAAGAAACAGAAATCTCATAGTAAATTTCTCC 620
 |||||
 Db 551 AAGTGCCAGAGAGCTTCAAGTCTTGAAGAAACAGAAATCTCATAGTAAATTTCTCC 610
 |||||
 QY 621 ATGCTCAACATACGGAAGAGTGTGCTCCCTTGGAGAGTCTCTGAAATTTCTTAAC 680
 |||||
 Db 611 ATGCTCAACATACGGAAGAGTGTGCTCCCTTGGAGAGTCTCTGAAATTTCTTAAC 670
 |||||
 QY 681 CTTGATCTCTTGGCTTCATGCTGACAGAGAGTCTAGATGAGAGCTTGGATAGAGGG 740
 |||||
 Db 671 CTTGATCTCTTGGCTTCATGCTGACAGAGAGTCTAGATGAGAGCTTGGATAGAGGG 730
 |||||
 QY 741 TGGATGACATGCTTGTCTGCTGCGGCTCTTGGAGAACTAAGAGATTTTCAAGAGCTATA 800
 |||||
 Db 731 TGGATGACATGCTTGTCTGCTGCGGCTCTTGGAGAACTAAGAGATTTTCAAGAGCTATA 790
 |||||
 QY 801 ATCAGAAATGTTTCGGAATAATGCCAGAGCTATCAACATGATTTTCCATCAATTCG 860
 |||||
 Db 791 ATCAGAAATGTTTCGGAATAATGCCAGAGCTATCAACATGATTTTCCATCAATTCG 850
 |||||
 QY 861 GCTTCAAGAAATTTTCAAGAGTACTGATCACTGAGGAAATGCACTGAGACTAGTA 920
 |||||
 Db 851 GCTTCAAGAAATTTTCAAGAGTACTGATCACTGAGGAAATGCACTGAGACTAGTA 910
 |||||
 QY 921 ACCAGCTTCTAAGAAAG-----GACCTGATCCCATTTGCCCTG 938
 |||||
 Db 911 ACCAGCTTCTAAGAAAGTTTAGGCTCTGAACAAGTACTAAGAGATATTCGCCGGA 970
 |||||
 QY 939 -----GACCTGATCCCATTTGCCCTG 962
 |||||
 Db 971 AACAAACCGATGGGTTAAAAACGTTTGTGACACACCTGGTCCCATTTGCCCTG 1030
 |||||
 QY 963 TCTATGCTTAAAGAGTATCTGATGTCGAAAGTGGAGAGTCTGTTCTTGAACCTGCTC 1022
 |||||
 Db 1031 TCTATGCTTAAAGAGTATCTGATGTCGAAAGTGGAGAGTCTGTTCTTGAACCTGCTC 1090
 |||||
 QY 1023 TTGAAATCGTCAAGAGTTTATCCAGGCGCAAGGCTTACAGCCATCAATTAAGATGC 1082
 |||||
 Db 1091 TTGAAATCGTCAAGAGTTTATCCAGGCGCAAGGCTTACAGCCATCAATTAAGATGC 1150
 |||||
 QY 1083 CATCAATGAAAGCTGAGAACAGAGAGTATATCACTGTGTGACCTGTGTGATGAATCA 1142
 |||||
 Db 1151 CATCAATGAAAGCTGAGAACAGAGAGTATATCACTGTGTGACCTGTGTGATGAATCA 1210
 |||||
 QY 1143 TCATTGGGATCGGAATGGGACGCAATAAATCCAAATCCACTTGAACCACTGA 1202
 |||||
 Db 1211 TCATTGGGATCGGAATGGGACGCAATAAATCCAAATCCACTTGAACCACTGA 1270
 |||||
 QY 1203 AGAAAAAGAGATTTGAGTCAATGCTGTCAACACATAGAAAGTCAAGGTCTTCCC 1262
 |||||
 Db 1271 AGAAAAAGAGATTTGAGTCAATGCTGTCAACACATAGAAAGTCAAGGTCTTCCC 1330
 |||||
 QY 1263 CAGACTTAACAAAGAACTAAAGAGAGGATCCCGAGGAGCAAAATGATCAAGCTGA 1322
 |||||
 Db 1331 CAGACTTAACAAAGAACTAAAGAGAGGATCCCGAGGAGCAAAATGATCAAGCTGA 1390
 |||||
 QY 1323 AATGACGCTTTAAGACATGTCAGTGGCTTTGAAAAGTGTGGGATCCAGTTCA 1382
 |||||
 Db 1391 AATGACGCTTTAAGACATGTCAGTGGCTTTGAAAAGTGTGGGATCCAGTTCA 1450
 |||||
 QY 1383 GGAAGGAGGGTATGTTGTCTCCAGTCTGGGAGAAAGAGTGTATGCGGAATTTCTG 1442
 |||||
 Db 1451 GGAAGGAGGGTATGTTGTCTCCAGTCTGGGAGAAAGAGTGTATGCGGAATTTCTG 1510
 |||||
 QY 1443 CATAGCAGAAAAGCTCCACCATTTCTTTGATGATGATTTAAAGTCTCAGTTCTTA 1502
 |||||
 Db 1511 CATAGCAGAAAAGCTCCACCATTTCTTTGATGATGATTTAAAGTCTCAGTTCTTA 1570
 |||||
 QY 1503 TAAATGAAACAGCAGGCTTGTCAAGTCTTGTGCTGATGTTGTCTGAATGATGTA 1562
 |||||

OTHER INFORMATION: Incyte ID No: 977391.2
NAME/KEY: unsure
LOCATION: 2106
OTHER INFORMATION: a, t, c, g, or other
US-60-324-185-33690

Query Match 93.5%; Score 1909; DB 76; Length 2109;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 2022; Conservative 0; Mismatches 6; Indels 80; Gaps 3;

QY 6 ATAAGATGGCGTCCGT -GGCGGCTGCACGAGCAGTTCGTGTGGGAGTGGGCTCAGAGGC 64
DB 3 AAAAGATGGCGTCCGTGGGCGGCTGCACGAGCAGTTCGTGTGGGAGTGGGCTCAGAGGC 62
QY 65 CTGCAACGAGACCTTACTCTTGTAGTGTCTGTGGGGCCACGGGGCACCGGCAATTCACG 124
DB 63 CTGCAACGAGACCTTACTCTTGTAGTGTCTGTGGGGCCACGGGGCACCGGCAATTCACG 122
QY 125 CTGGCGTGTGAGTGGCCAGGCGGCTGGGGGGTGAAGTGTGAGCGCTGACCTGCAATGCA 184
DB 123 CTGGCGTGTGAGTGGCCAGGCGGCTGGGGGGTGAAGTGTGAGCGCTGACCTGCAATGCA 182
QY 185 GTCTATGAAGGCTTAGACATCATCACCAACAGGTTTCTGCCAAGAGCAGAGATCTGC 244
DB 183 GTCTATGAAGGCTTAGACATCATCACCAACAGGTTTCTGCCAAGAGCAGAGATCTGC 242
QY 245 CGGCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 304
DB 243 CGGCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 302
QY 305 AGAATAGAGCAACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 364
DB 303 AGAATAGAGCAACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 362
QY 365 GTGGGAGAGCAATATTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 424
DB 363 GTGGGAGAGCAATATTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
QY 425 CCCGAGAGATGGGCACTGAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 484
DB 423 CCCGAGAGATGGGCACTGAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 482
QY 485 GGTCTGTACTTACAAAGCCCTAAGCGAGGTGGACCCAGCAAAATGGTGGCAAGTGCAT 544
DB 483 GGTCTGTACTTACAAAGCCCTAAGCGAGGTGGACCCAGCAAAATGGTGGCAAGTGCAT 542
QY 545 CCACATGACAAAGCAAGTGGCCAGAGGCTTCAAGTTTGTGAAGAAACAGGATCTCT 604
DB 543 CCACATGACAAAGCAAGTGGCCAGAGGCTTCAAGTTTGTGAAGAAACAGGATCTCT 602
QY 605 CATAGTGAATTTCTCATGCTCAACATPAGGAAAGTGGTGGTCCCTTGGAGTCTT 664
DB 603 CATAGTGAATTTCTCATGCTCAACATPAGGAAAGTGGTGGTCCCTTGGAGTCTT 662
QY 665 CTGAAGTTCCTTAACCTTGCATCTTGGCTCATGCTGACAGGAGGAGTCTAGATGAG 724
DB 663 CTGAAGTTCCTTAACCTTGCATCTTGGCTCATGCTGACAGGAGGAGTCTAGATGAG 722
QY 725 CGCTTGATTAAGAGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 784
DB 723 CGCTTGATTAAGAGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 782
QY 785 TTTTCACAGAGCTATTAATCAGAAATGTTTGGGAAATAGCCAGGACTATCAACATGCT 844
DB 783 TTTTCACAGAGCTATTAATCAGAAATGTTTGGGAAATAGCCAGGACTATCAACATGCT 842
QY 845 ATCTTCAATCAATGGCTTCAAGAAATTTCAAGAGTATTCAGAGTATCTGAGGAAATGTC 904
DB 843 ATCTTCAATCAATGGCTTCAAGAAATTTCAAGAGTATTCAGAGTATCTGAGGAAATGTC 902
QY 905 ACACGTGAGAGCTAGTAACAGAGCTTCTAAAGAAAG----- 938
DB 903 ACACGTGAGAGCTAGTAACAGAGCTTCTAAAGAAAGGTAATTGAGGCTCTGAAACAAAGTAAT 902

QY 939 -----GACCTGCT 946
DB 963 AAGAGATATGCCCGGAACAAAACGATGGGTTAAAAACCGTTTGTGACAGACCTGCT 1022
QY 947 CCCATTTGCCCGCTGCTTATAGGCTTATAGAGTATCTGATGATCTGAAAGTGGAGAGTCT 1006
DB 1023 CCCATTTGCCCGCTGCTTATAGGCTTATAGAGTATCTGATGATCTGAAAGTGGAGAGTCT 1082
QY 1007 GTTCTTGAACCTGCTCTTGAATCTGTCAAAAGTTTATCCAGGGCCACAAGCTTACAGCC 1066
DB 1083 GTTCTTGAACCTGCTCTTGAATCTGTCAAAAGTTTATCCAGGGCCACAAGCTTACAGCC 1142
QY 1067 ACTCCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1126
DB 1143 ACTCCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1202
QY 1127 CTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1186
DB 1203 CTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1262
QY 1187 CACTTGAACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1246
DB 1263 CACTTGAACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1322
QY 1247 AGTCAGAGTGTTCCTCCAGACTTAAACAAAGAACTTAAAGGAGGATCCCGAGGCGAG 1306
DB 1323 AGTCAGAGTGTTCCTCCAGACTTAAACAAAGAACTTAAAGGAGGATCCCGAGGCGAG 1382
QY 1307 AATGATCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366
DB 1383 AATGATCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1442
QY 1367 GTGGGATCCAGTTCAGAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1426
DB 1443 GTGGGATCCAGTTCAGAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1502
QY 1427 TATGCGGAATTCCTGCTATGAGCAAGAAAGTCCACATTTTCTTGAACCTTAAAGTCTTA 1486
DB 1503 TATGCGGAATTCCTGCTATGAGCAAGAAAGTCCACATTTTCTTGAACCTTAAAGTCTTA 1562
QY 1487 AGCTCAGCTTCTCTAATATAGAAACAGCAGGCTTGTGACGCTCTTGTGGTGGTATG 1546
DB 1563 AGCTCAGCTTCTCTAATATAGAAACAGCAGGCTTGTGACGCTCTTGTGGTGGTATG 1622
QY 1547 GTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1606
DB 1623 GTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1682
QY 1607 TTAATTAAGACACAGATTCACATTTTATACATGAGAGATCTTGTGGTGAATA 1666
DB 1683 TTAATTAAGACACAGATTCACATTTTATACATGAGAGATCTTGTGGTGAATA 1742
QY 1667 CCAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1726
DB 1743 CCAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1801
QY 1727 CTAATTTCCAGATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1786
DB 1802 CTAATTTCCAGATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1861
QY 1787 TAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1846
DB 1862 TAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1921
QY 1847 TAGTCTCGGAATTTACAGAGAGGAGGAAATCACTGAGAACTGTGACATGAGAC 1906
DB 1922 TAGTCTCGGAATTTACAGAGAGGAGGAAATCACTGAGAACTGTGACATGAGAC 1981
QY 1907 TTAAGAGCAAAAGACTTTGAAATTTGCAAGCTGCTCATGTGTGATTAATATCATGCTG 1966
DB 1982 TTAAGAGCAAAAGACTTTGAAATTTGCAAGCTGCTCATGTGTGATTAATATCATGCTG 2041

QY	301	CTTCAGAAATAGAGCAACTCTCTGATATGAGATATATTTGGCCAGCAAAATTCAT	360
Db	301	CTTCAGAAATAGAGCAACTCTCTGATATGAGATATATTTGGCCAGCAAAATTCAT	360
QY	361	TGTTGTGGAGAGCAAAATATTACATTTGAATCTCTGCTTGGAAGTCTTCTCAATAC	420
Db	361	TGTTGTGGAGAGCAAAATATTACATTTGAATCTCTGCTTGGAAGTCTTCTCAATAC	420
QY	421	CAAGCCCCAGAGATGGGCACTGAGAAAGTATTGACGAAAAGTGGAGCTTGAAGA	480
Db	421	CAAGCCCCAGAGATGGGCACTGAGAAAGTATTGACGAAAAGTGGAGCTTGAAGA	480
QY	481	GGATGATCTTGTACTTCACAAAAGCCTTAAGCCAGGGTGGACCCAGAAATGGCTGCCAACT	540
Db	481	GGATGATCTTGTACTTCACAAAAGCCTTAAGCCAGGGTGGACCCAGAAATGGCTGCCAACT	540
QY	541	GCATCCACATAGACAAAGCGCAAAAGTGGCCAGAGACTTGCAGTTTTTGAAGAACAAGAAAT	600
Db	541	GCATCCACATAGACAAAGCGCAAAAGTGGCCAGAGACTTGCAGTTTTTGAAGAACAAGAAAT	600
QY	601	CTCTCATAGTGAATTTCTCATATGTCACATACGGAAGAAGTGGTGGTCCCTTGGAGG	660
Db	601	CTCTCATAGTGAATTTCTCATATGTCACATACGGAAGAAGTGGTGGTCCCTTGGAGG	660
QY	661	TCCCTGAAAGTCTCTAAACCCTTGCATTCCTTTGGGCTTACGGTGAACAGCAAGTCTAGA	720
Db	661	TCCCTGAAAGTCTCTAAACCCTTGCATTCCTTTGGGCTTACGGTGAACAGCAAGTCTAGA	720
QY	721	TGAGCGCTTGGATPAAGAGGGTGGATGACATGCTTGGCTGGGCTCTTGGAGAACTAAG	780
Db	721	TGAGCGCTTGGATPAAGAGGGTGGATGACATGCTTGGCTGGGCTCTTGGAGAACTAAG	780
QY	781	AGATTTTCACAGACGCTATATATAGAGAAGTGTTCGGAATATAGCCAGACTATCAACA	840
Db	781	AGATTTTCACAGACGCTATATATAGAGAAGTGTTCGGAATATAGCCAGACTATCAACA	840
QY	841	TGATATCTTCCAAATCAATTTGGCTTCAAGAAATTCACAGATCTCGATACCTGAGAGAA	900
Db	841	TGATATCTTCCAAATCAATTTGGCTTCAAGAAATTCACAGATCTCGATACCTGAGAGAA	900
QY	901	ATGCACACTGGAGACTAGTAACCAAGCTTCTAAAGAAAG-----	938
Db	901	ATGCACACTGGAGACTAGTAACCAAGCTTCTAAAGAAAG-----	938
QY	939	-----GACC 942	
Db	961	AACATAAGAGATATGCCCGGAACAAAACCGATGGTCTTAAAAACCGTTTTTGGACGACCC	1020
QY	961	AACATAAGAGATATGCCCGGAACAAAACCGATGGTCTTAAAAACCGTTTTTGGACGACCC	1020
QY	961	TGGTCCCATTTGTCCCCCTCTATATGGCTTAAAGAGTATCTGATGTCTCGAATGGGAGAGA	1002
Db	1021	TGGTCCCATTTGTCCCCCTCTATATGGCTTAAAGAGTATCTGATGTCTCGAATGGGAGAGA	1002
QY	1003	GTCGTGCTTAAACCTGCTCTGTGAAGATCGGCAAACTTTCATCCAGGGGCAACAACCTAC	1062
Db	1081	GTCGTGCTTAAACCTGCTCTGTGAAGATCGGCAAACTTTCATCCAGGGGCAACAACCTAC	1140
QY	1063	AGCCACTCCAAATAAGATGCCATACATGAAGCTTGAGAAACAAGAGATTAACCTGTG	1122
Db	1141	AGCCACTCCAAATAAGATGCCATACATGAAGCTTGAGAAACAAGAGATTAACCTGTG	1200
QY	1123	TGACCTCTGTGATGCAATCATATGGGGATGCGCAATGGGCAAGCGCCACATATAAATCCA	1182
Db	1201	TGACCTCTGTGATGCAATCATATGGGGATGCGCAATGGGCAAGCGCCACATATAAATCCA	1260
QY	1183	ATCCCACTTGAACCACTGGAAGAAAGAAAGAAATTTGAGACTGAGATGGCTGTCAACACAT	1242
Db	1261	ATCCCACTTGAACCACTGGAAGAAAGAAAGAAATTTGAGACTGAGATGGCTGTCAACACAT	1320
QY	1243	AGAAAGTCAGAGTCTTTCGCCAGACTTAAACAAGAAGCTTAAAGGAGAGGATCCCCACAG	1302
Db	1321	AGAAAGTCAGAGTCTTTCGCCAGACTTAAACAAGAAGCTTAAAGGAGAGGATCCCCACAG	1380
QY	1303	GCAGATGATATCAGAGCTGAATATGCAAGCTTTAAGAGACATGTCCAGTGGCTTTGGAAA	1362

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

OY 721 TGAGCGCTTGATAGAGGGTGATGACATGCTTGCTGCTGGGCTCTGGAGGAGCAATAG 780
    |||||||
Db 725 TGAGCGCTTGATAGAGGGTGATGACATGCTTGCTGCTGGGCTCTGGAGGAGCAATAG 784
OY 781 AGATTTTCACAGAGCGCTATATGAGAAAGATTTTCGAAAGATAGCAGAGCATATCACA 840
    |||||||
Db 785 AGATTTTCACAGAGCGCTATATGAGAAAGATTTTCGAAAGATAGCAGAGCATATCACA 844
OY 841 TGGTATCTTCCAAATGATGGCTTCAAGGAATTTCAAGATACCTGATCACTAGAGGAAA 900
    |||||||
Db 845 TGGTATCTTCCAAATGATGGCTTCAAGGAATTTCAAGATACCTGATCACTAGAGGAAA 904
OY 901 ATGCACACTGAGAGACTAGTAACAGCTTCTAAGAAAG----- 938
    |||||||
Db 905 ATGCACACTGAGAGACTAGTAACAGCTTCTAAGAAAGAGTATTGAGGCTCTGAACAGT 964
OY 939 -----GACC 942
    ||||
Db 965 AACTAAGAGATATGCCGGAACAAACCGATGGTTAAACCGTTTGTGAGCAGACC 1024
OY 943 TGGTCCCATTTGCCCTGCTATGAGGTATGATGCTCGAAGTGGGAGGA 1002
    |||||||
Db 1025 TGGTCCCATTTGCCCTGCTATGAGGTATGATGCTCGAAGTGGGAGGA 1084
OY 1003 GTCTGTCTTGAACTGCTCTTGAATCGTGCAGAACTTTCATCCAGGCGCACAAGCTAC 1062
    |||||||
Db 1085 GTCTGTCTTGAACTGCTCTTGAATCGTGCAGAACTTTCATCCAGGCGCACAAGCTAC 1144
OY 1063 AGCCACTCCAAATAGATGCAATGCAATGAGCTGAGAACAGAGAATATACCTGTG 1122
    |||||||
Db 1145 AGCCACTCCAAATAGATGCAATGCAATGAGCTGAGAACAGAGAATATATACCTGTG 1204
OY 1123 TGAACCTCTGTGATGCAATCATCTTGGGATCGCGAATGGGCGCACATATAATCCAA 1182
    |||||||
Db 1205 TGAACCTCTGTGATGCAATCATCTTGGGATCGCGAATGGGCGCACATATAATCCAA 1264
OY 1183 ATCCCACTTGAAACCACTGAGAAAGAAAGAAATTTGATGATGCTCTCAACACAT 1242
    |||||||
Db 1265 ATCCCACTTGAAACCACTGAGAAAGAAAGAAATTTGATGATGCTCTCAACACAT 1324
OY 1243 AGAAAGTCAGAGTGTTCCTCCAGACTATTAACAAGAACTTAAGAGGAGGATCCGAG 1302
    |||||||
Db 1325 AGAAAGTCAGAGTGTTCCTCCAGACTATTAACAAGAACTTAAGAGAGGATCCGAG 1384
OY 1303 GCAGAAATGATCAAGAGCTGAAATGCAAGCTTTAAGAGACATGTCAGTGGCTTTGAAA 1362
    |||||||
Db 1385 GCAGAAATGATCAAGAGCTGAAATGCAAGCTTTAAGAGACATGTCAGTGGCTTTGAAA 1444
OY 1363 GGTGGTGGGGATGATGATGAGAGGAGGAGGTATGTTGTCCTCCAGTCTGGGCAAGGA 1422
    |||||||
Db 1445 GGTGGTGGGGATGATGATGAGAGGAGGAGGTATGTTGTCCTCCAGTCTGGGCAAGGA 1504
OY 1423 GTGCTATGGGGAATCTCTGATAGAGAAAGCTCCACCATTTTCTTTGATGTGTT 1482
    |||||||
Db 1505 GTGCTATGGGGAATCTCTGATAGAGAAAGCTCCACCATTTTCTTTGATGTGTT 1564
OY 1483 TTTAAATCTCAGCTTCTATATATAGAAACAGAGCTTTGTGAGCTCTGTGTGCTG 1542
    |||||||
Db 1565 TTTAAATCTCAGCTTCTATATATAGAAACAGAGCTTTGTGAGCTCTGTGTGCTG 1624
OY 1543 ATGTGCTGGAATGATGATGATGAGAAAGCAATTTTCTTTGAAACCTTAAGGT 1602
    |||||||
Db 1625 ATGTGCTGGAATGATGATGATGAGAAAGCAATTTTCTTTGAAACCTTAAGGT 1684
OY 1603 TCTATTATTAAGAGAGAGAGATTCACATTTTATACATGAGATCTTCTTTGTGTG 1662
    |||||||
Db 1685 TCTATTATTAAGAGAGAGAGATTCACATTTTATACATGAGATCTTCTTTGTGTG 1744
OY 1663 AATACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1722
    |||||||
Db 1745 AATACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1803

```

```

OY 1723 TTAATCAATTCAGAGTCTTTTGTAGATGACTGAAAGTATTTGTGAGCCACATATGGGA 1782
    |||||||
Db 1804 TTAATCAATTCAGAGTCTTTTGTAGATGACTGAAAGTATTTGTGAGCCACATATGGGA 1863
OY 1783 GTTCTAGATTTGAGTAATGGCAGAGAAAGCCATCTCATTTGATGATGATTAAGTAAGC 1842
    |||||||
Db 1864 GTTCTAGATTTGAGTAATGGCAGAGAAAGCCATCTCATTTGATGATGATTAAGTAAGC 1923
OY 1843 AAATAGTCTCGGAATTTCTACAGAGAGAGGGAATCAAGACTGAGGAAGCTGTGACATA 1902
    |||||||
Db 1924 AAATAGTCTCGGAATTTCTACAGAGAGAGGGAATCAAGACTGAGGAAGCTGTGACATA 1983
OY 1903 GGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1962
    |||||||
Db 1984 GGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2043
OY 1963 GCTGTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2022
    |||||||
Db 2044 GCTGTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2103
OY 2023 TTACAAGAAAAA 2041
    ||||
Db 2104 TTAAGAAAAA 2122

```

```

RESULT 6
US-10-133-013-165
; Sequence 165, Application US/10133013
; GENERAL INFORMATION:
; APPLICANT: Astromoff, Anna
; APPLICANT: Bandman, Olga
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES ASSOCIATED WITH VASCULAR DISEASE
; FILE REFERENCE: PA-0049 US
; CURRENT APPLICATION NUMBER: US/10/133, 013
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/287, 067
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 271
; SOFTWARE: PERL Program
; SEQ ID NO: 165
; LENGTH: 2108
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 977391.15
US-10-133-013-165

```

```

Query Match 94.2%; Score 1923.6; DB 40; Length 2108;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 2025; Conservative 0; Mismatches 4; Indels 79; Gaps 2;
OY 1 CTGCCATTAAGATGCGTCCGTGCGGCTGCGAGCAGAGCATGCTGCGGAGTGGCTCAG 60
    |||||||
Db 1 CTGCCATTAAGATGCGTCCGTGCGGCTGCGAGCAGAGCATGCTGCGGAGTGGCTCAG 60
OY 61 GGGCTGCAAGGAGACCTTACTTGTAGTATCTTCGCGGGCCACGCGCAAGCAATC 120
    |||||||
Db 61 GGGCTGCAAGGAGACCTTACTTGTAGTATCTTCGCGGGCCACGCGCAAGCAATC 120
OY 121 CAGCGTGGGCTTACAGTGGGCGGAGGCGTGGGCGGTGAGATGTCAGGCTGACATC 180
    |||||||
Db 121 CAGCGTGGGCTTACAGTGGGCGGAGGCGTGGGCGGTGAGATGTCAGGCTGACATC 180
OY 181 GCAGGTCTATGAAGGCTAGACATCAACCAAGAGGTTTTCGCCAAGAGAGAGAAAT 240
    |||||||
Db 181 GCAGGTCTATGAAGGCTAGACATCAACCAAGAGGTTTTCGCCAAGAGAGAGAAAT 240
OY 241 CTGCCGACACACATATACGCTTTGTGATCTCTTGTGACCAATTAACAGTGTGGA 300
    |||||||
Db 241 CTGCCGACACACATATACGCTTTGTGATCTCTTGTGACCAATTAACAGTGTGGA 300

```

	Prior Application Number:	60/276,025
	Prior Filing Date:	2001-03-14
	Prior Application Number:	60/325,149
	Prior Filing Date:	2001-09-26
	Prior Application Number:	60/276,026
	Prior Filing Date:	2001-03-14
	Prior Application Number:	60/324,967
	Prior Filing Date:	2001/09/26
	Prior Application Number:	60/311,732
	Prior Filing Date:	2001-08-10
	Prior Application Number:	60/325,102
	Prior Filing Date:	2001-09-26
	Prior Application Number:	60/323,580
	Prior Filing Date:	2001-09-19
	Number of Seq ID NOS:	363
	Software:	FastSeq for Windows Version 4.0
	Seq ID NO:	150
	Length:	2129
	Type:	DNA
	Organism:	Homo sapiens
	US-10-097-340-150	
	Query Match	94.5%; Score 1929.4; DB 39; Length 2129;
	Best Local Similarity	95.8%; Pred. No. 0;
	Matches 2029; Conservative	5; Mismatches 6; Indels 79; Gaps 2.
OY	1 CTGCATTAAGATGGCGTCCGTGCGGCTGCACAGCAGACTTCCTGTGGGAGAGGGCTCAG	60
Db	5 CTGCCATTAGAAGGGGTCCGTGCGGCTGCACAGCAGAGTTCTGTGGGAGATGGCTCAG	64
OY	61 GGCGCTGCAAGGACCCTCACTCTTGTTAGTAGTTCTGGGGCCACAGGGGCAACGAATC	120
Db	65 GGCGCTGCAAGGACCCTCACTCTTGTTAGTAGTTCTGGGGCCACAGGGGCAACGAATC	124
OY	121 CACGCTGGCGTTGACAGTAAAGCCACGCGCTGGGGGTGAGATCGTACGCTCATCCAT	180
Db	125 CACGCTGGCGTTGACAGTAAAGCCACGCGCTGGGGGTGAGATCGTACGCTCATCCAT	184
OY	181 GCAGGTCTAATGAGGCTCTAGATCATCATCAACAAGTTTGGCCCAAGAGCAGAGAAAT	240
Db	185 GCAGGTCTAATGAGGCTCTAGATCATCATCAACAAGTTTGGCCCAAGAGCAGAGAAAT	244
OY	241 CTCGCGGCAACCATGATAGCTTTGTGATCCTCTTGTGACCAATTACACAGTGTGGA	300
Db	245 CTCGCGGCAACCATGATAGCTTTGTGATCCTCTTGTGACCAATTACACAGTGTGGA	304
OY	301 CTTCAAAATAGAGCAACTGCTGTGATTAAGATATATTTGCCGAGACAAAATTCCTAT	360
Db	305 CTTCAAAATAGAGCAACTGCTGTGATTAAGATATATTTGCCGAGACAAAATTCCTAT	364
OY	361 TGTGTGGAGAGAACCAATTATTACTTTAAATCTCTGCTCTGGAAGTCTTGTCAATAC	420
Db	365 TGTGTGGAGAGAACCAATTATTACTTTAAATCTCTGCTCTGGAAGTCTTGTCAATAC	424
OY	421 CAAGCCCCGAGGATGGGCACTGAGAAAGTATGACCAAAAAAGTGGAGCTTGAAGAAG	480
Db	425 CAAGCCCCGAGGATGGGCACTGAGAAAGTATGACCAAAAAAGTGGAGCTTGAAGAAG	484
OY	481 GGATGTCCTTGTACTTCACAAAAGCCTAAGCCAGGTGAGCCCGAAGTGGCTCCAACT	540
Db	485 GGATGTCCTTGTACTTCACAAAAGCCTAAGCCAGGTGAGCCCGAAGTGGCTCCAACT	544
OY	541 GCATCACATGACAAAAGGAGTGGCCAGAGCTTGCAGAGTTTTTGAAAGAACAGAAAT	600
Db	545 GCATCACATGACAAAAGGAGTGGCCAGAGCTTGCAGAGTTTTTGAAAGAACAGAAAT	604
OY	601 CTCCTAATAGTAATTTCTCATCGTGAACATACGGAAGAAAGTGGTGGTCCCTTGAGG	660
Db	605 CTCCTAATAGTAATTTCTCATCGTGAACATACGGAAGAAAGTGGTGGTCCCTTGAGG	664
OY	661 TCCTCGAAGTTCCTCAACCTTGACATCTTGGCTTCATGCTGACAGGACAGTTCTAGA	720
Db	665 TCCTCGAAGTTCCTCAACCTTGACATCTTGGCTTCATGCTGACAGGACAGTTCTAGA	724

PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 152
LENGTH: 2129
TYPE: DNA
ORGANISM: Homo sapiens
US-10-097-340-152

Query Match 94.6%; Score 1931; DB 39; Length 2129;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 2033; Conservative 1; Mismatches 6; Indels 79; Gaps 2;

QY 1 CTGCCATAGATGGCGTCCGTGCGGCTGCAGAGCAGTTCTGTGGGAGTGGGCTCAG 60
DB 5 CTGCCATAGATGGCGTCCGTGCGGCTGCAGAGCAGTTCTGTGGGAGTGGGCTCAG 64

QY 61 GGGCCCTGCAAGCAACCTTACTCTGTAGTGAATTTCTGGGGCCACGGGCAACATC 120
DB 65 GGGCCCTGCAAGCAACCTTACTCTGTAGTGAATTTCTGGGGCCACGGGCAACATC 124

QY 121 CAGCGTGGCTGAGCTAGAGCCAGCGGCTGGGGGTGAGATGTGCTGAGCTGACTCAT 180
DB 125 CAGCGTGGCTGAGCTAGAGCCAGCGGCTGGGGGTGAGATGTGCTGAGCTGACTCAT 184

QY 181 GCAGGCTGTAGAGGCTTAGACATCATCACCAACAGGTTTCTGCCAGAGAGAGAT 240
DB 185 GCAGGCTGTAGAGGCTTAGACATCATCACCAACAGGTTTCTGCCAGAGAGAGAT 244

QY 241 CTGCGCGGACACATGATGATGCTTTGTGATCTCTTGTGACCAATTACACAGTGTGA 300
DB 245 CTGCGCGGACACATGATGATGCTTTGTGATCTCTTGTGACCAATTACACAGTGTGA 304

QY 301 CTTCAGAAATAGAGCACTGCTGATTTGAATTTGGCCGAGACAAATTCCTAT 360
DB 305 CTTCAGAAATAGAGCACTGCTGATTTGAATTTGGCCGAGACAAATTCCTAT 364

QY 361 TGTGTGGAGAGAACCAATTATTACATTTGATCTGCTGTGAAAGTTCTTGTCAATAC 420
DB 365 TGTGTGGAGAGAACCAATTATTACATTTGATCTGCTGTGAAAGTTCTTGTCAATAC 424

QY 421 CAGGCCCGAGAGAGGCGCTAGAGAGAGTGAACCCGAAAGTGGAGCTTGAAGAAGA 480
DB 425 CAGGCCCGAGAGAGGCGCTAGAGAGAGTGAACCCGAAAGTGGAGCTTGAAGAAGA 484

QY 481 GGATGCTTGTACTTCAACCAAGCCCTAAGCCAGGTGAGCCAGAAATGCTGCCAAGCT 540
DB 485 GGATGCTTGTACTTCAACCAAGCCCTAAGCCAGGTGAGCCAGAAATGCTGCCAAGCT 544

QY 541 GCATTCACATGACAAACGAAAGTGGCGAGAGCTTGCAAGTTTGAAGAAACAGGAAT 600
DB 545 GCATTCACATGACAAACGAAAGTGGCGAGAGCTTGCAAGTTTGAAGAAACAGGAAT 604

QY 601 CTCTCATAGGAATTTCTCATGCTCAACATAGAGAAAGAGTGTGCTCCCTTGAGAG 660
DB 605 CTCTCATAGGAATTTCTCATGCTCAACATAGAGAAAGAGTGTGCTCCCTTGAGAG 664

QY 661 TCCCTGAAATTTCTTAACCTTGATCTTGGCTCATGCTGACAGGAGAGTTCAGA 720
DB 665 TCCCTGAAATTTCTTAACCTTGATCTTGGCTCATGCTGACAGGAGAGTTCAGA 724

QY 721 TGAGCGCTGGATTAAGAGGTGATGATGCTTGTGCTGAGGCTCTTGGAGAACTAAG 780
DB 725 TGAGCGCTGGATTAAGAGGTGATGATGCTTGTGCTGAGGCTCTTGGAGAACTAAG 784

QY 781 AGATTTTCAAGAGCGTATATAGAAAGATGTTGGAGAAATAGCCAGACATATACAA 840
DB 785 AGATTTTCAAGAGCGTATATAGAAAGATGTTGGAGAAATAGCCAGACATATACAA 844

QY 841 TGGATCTTCCATCAATTTGGCTTCAAGAAATTTCAAGAGTACCTGATCAGAGGAAA 900
DB 845 TGGATCTTCCATCAATTTGGCTTCAAGAAATTTCAAGAGTACCTGATCAGAGGAAA 904

QY 901 ATGCACACTGGAGACTAGTAACCACTTCTAAGAAAG----- 938
DB 905 ATGCACACTGGAGACTAGTAACCACTTCTAAGAAAGATTTGAGGCTGTGAACAAAGT 964

QY 939 -----GACC 942
DB 965 AACTAAGATATGCCCCGAAACAAACCGATGGGTTAAACCGTTTGTGAGCACACC 1024

QY 943 TGGTCCATGTGCCCCCTGCTATGAGCTTAGAGGTATGATGTGCGAAGTGGGAGGA 1002
DB 1025 TGGTCCATGTGCCCCCTGCTATGAGCTTAGAGGTATGATGTGCGAAGTGGGAGGA 1084

QY 1003 GTCTGTTCTTGAACCTGCTTTGAATCGTCAAAAGTTTATCCAGGCGCACAGCTTAC 1062
DB 1085 GTCTGTTCTTGAACCTGCTTTGAATCGTCAAAAGTTTATCCAGGCGCACAGCTTAC 1144

QY 1063 AGCCACTCCAAATAAGATGCCATACATGAGCTGAGAACCAAGAAAGTTATCACTGTG 1122
DB 1145 AGCCACTCCAAATAAGATGCCATACATGAGCTGAGAACCAAGAAAGTTATCACTGTG 1204

QY 1123 TGAACCTGTGATGCAATCATCATTTGGGGATCGGAAATGGGCGACATAAATCCAA 1182
DB 1205 TGAACCTGTGATGCAATCATCATTTGGGGATCGGAAATGGGCGACATAAATCCAA 1264

QY 1183 ATCCACTTGAACCACTGAGAAAGAAAGAAAGTTGAGCTCAGATGCTGTCAACACAT 1242
DB 1265 ATCCACTTGAACCACTGAGAAAGAAAGAAAGTTGAGCTCAGATGCTGTCAACACAT 1324

QY 1243 AGAAAGTCAAGTGTGTTCCCGAGACTATTAACAAGAACTTAAGGAGAGGATCCCAAG 1302
DB 1325 AGAAAGTCAAGTGTGTTCCCGAGACTATTAACAAGAACTTAAGGAGAGGATCCCAAG 1384

QY 1303 GCAGAAATGATCAAGAGCTGAATGCAGCTTTAAGACATGTCACATGGCTTTGGAAA 1362
DB 1385 GCAGAAATGATCAAGAGCTGAATGCAGCTTTAAGACATGTCACATGGCTTTGGAAA 1444

QY 1363 GGTGTGGGAGTCCAGTTTCAAGAGAGGAGTATGTTGTCTCCACTGTGGGCAAGA 1422
DB 1445 GGTGTGGGAGTCCAGTTTCAAGAGAGGAGTATGTTGTCTCCACTGTGGGCAAGA 1504

QY 1423 GTGCTATGCGGAATTTCTCTCATAGCAAGAAAGCTCCACATTTTGTGATGTGTT 1482
DB 1505 GTGCTATGCGGAATTTCTCTCATAGCAAGAAAGCTCCACATTTTGTGATGTGTT 1564

QY 1483 TTAAGTCTCAGCTTCTCTAATATAGAACAGAGCTTGTGAGCTCTTGTGTGCTG 1542
DB 1565 TTAAGTCTCAGCTTCTCTAATATAGAACAGAGCTTGTGAGCTCTTGTGTGCTG 1624

QY 1543 ATGTGTCTGGAATGATGTAGTTCAGGAAGCAATTTTCTTGTGAACCTTAAGGT 1602
DB 1625 ATGTGTCTGGAATGATGTAGTTCAGGAAGCAATTTTCTTGTGAACCTTAAGGT 1684

QY 1603 TCTATTAATTAAGCAAGCAAGATTCACATTTTATACATGAGAGATCTTGTGTG 1662
DB 1685 TCTATTAATTAAGCAAGCAAGATTCACATTTTATACATGAGAGATCTTGTGTG 1744

QY 1663 AATACAGAGATTAAGTCACTCCCTTTAAAGAAAGTTTATGCTCCCTGACTGCTGCTAAA 1722
DB 1745 AATACAGAGATTAAGTCACTCCCTTTAAAGAAAGTTTATGCTCCCTGACTGCTGCTAAA 1803

QY 1723 TTATCTAATTTCCAGATGCTTTTGTAGATGCTGAAGTATTTGTGAGCCACATATTGGGA 1782
DB 1804 TTATCTAATTTCCAGATGCTTTTGTAGATGCTGAAGTATTTGTGAGCCACATATTGGGA 1863

QY 1783 GTTCTGATTTGAGTGAATGAGGAGAAAGGCGCATCTCCATGAGATTAAGTGAACC 1842
DB 1864 GTTCTGATTTGAGTGAATGAGGAGAAAGGCGCATCTCCATGAGATTAAGTGAACC 1923

QY 1843 AAACATGTTCTCGGAATTTCAACAGAGAGAGGAATCAACTGAGAGAGCTGTGACATA 1902
DB 1924 AAACATGTTCTCGGAATTTCAACAGAGAGAGGAATCAACTGAGAGAGCTGTGACATA 1983

OY	541	GCATCCACATATACAAACGAAAGGGCGAGAGACTTGCAAGTTTTTGAAGAAACAGAGAT	600
Db	545	GCATCCACATATACAAACGAAAGGGCGAGAGACTTGCAAGTTTTTGAAGAAACAGAGAT	604
OY	601	CTCTCATAGTGAATTTCTCCATGCTCAACATACAGAAAGGTGGTGGTCCCTTGAGAG	660
Db	605	CTCTCATAGTGAATTTCTCCATGCTCAACATACAGAAAGGTGGTGGTCCCTTGAGAG	664
OY	661	TCCTCTGAAGTCTCTAACCCTTGACATCTTTGGCTTTCATGCTGACACGAGCTTCTAGA	720
Db	665	TCCTCTGAAGTCTCTAACCCTTGACATCTTTGGCTTTCATGCTGACACGAGCTTCTAGA	724
OY	721	TGAGGCGCTGGATTAAGAGGGGTGGATGACATGCTGCTGGTGGCTCTTGAGAGACTAAG	780
Db	725	TGAGGCGCTGGATTAAGAGGGGTGGATGACATGCTGCTGGTGGCTCTTGAGAGACTAAG	784
OY	781	AGATTTTTCACAGAGCGCTATATAGAAAGAAATGTTTGGAAAAATAGCAGAGACTATCAACA	840
Db	785	AGATTTTTCACAGAGCGCTATATAGAAAGAAATGTTTGGAAAAATAGCAGAGACTATCAACA	844
OY	841	TGCTATCTTCCATCAATCAATTGGCTTCAAGGAATTTACAGAGTACCTGATCATGAGGGAAA	900
Db	845	TGCTATCTTCCATCAATCAATTGGCTTCAAGGAATTTACAGAGTACCTGATCATGAGGGAAA	904
OY	901	ATGGACACTGGAGACTGTATTAACCAAGCTCTCTAAAGAAAG-----	938
Db	905	ATGGACACTGGAGACTGTATTAACCAAGCTCTCTAAAGAAAGATTAAGAGCTTGAAGAAAGT	964
OY	939	-----GACC	942
Db	945	ANCTAAGAGATATCCCGGGAAACAAAACGATGGCTTAATAAACCGTTTTTTAGCAGACAC	1024
OY	943	TGCTCCCATTTGTCCCCCTGTCTATATGGCTTAGAGATCTGATGTTCTCGAAGTGGGAGGA	1002
Db	1025	TGCTCCCATTTGTCCCCCTGTCTATATGGCTTAGAGATCTGATGTTCTCGAAGTGGGAGGA	1008
OY	1003	GTCGTCTCTTGAACCTGCTCTTGAATTCGTCGAAATTTTCATCCAGAGGCCACAAACCTTAC	1064
Db	1085	GTCGTCTCTTGAACCTGCTCTTGAATTCGTCGAAATTTTCATCCAGAGGCCACAAACCTTAC	1144
OY	1063	AGCCACTCCATAAAGATGGCATACATGACGTGAGAACAGAGAAAGTTATCACCCTGTG	1122
Db	1145	AGCCACTCCATAAAGATGGCATACATGACGTGAGAACAGAGAAAGTTATCACCCTGTG	1204
OY	1123	TGACCTCTGTGATCGAATCATCTTTGGGGATCGCGAATGGGAGCGCACATTAATTCOA	1182
Db	1205	TGACCTCTGTGATCGAATCATCTTTGGGGATCGCGAATGGGAGCGCACATTAATTCOA	1264
OY	1183	ATCCCACTTGAACCAACTGAAAGAAAAGAGAGATTGGACTAGATGCTGTACACCAT	1242
Db	1265	ATCCCACTTGAACCAACTGAAAGAAAAGAGAGATTGGACTAGATGCTGTACACCAT	1324
OY	1243	AGAAAGTCAGAGTCTTCCCAAGACATTAACAAAGAAAGCTTAAGAGGAAGGATCTCCCAAG	1302
Db	1325	AGAAAGTCAGAGTCTTCCCAAGACATTAACAAAGAAAGCTTAAGAGGAAGGATCTCCCAAG	1384
OY	1303	GCAGAAATGATCAAGAGCTGGAATGACAGCGTTTAAAGACATGTCAGAGTGGCTTTGGAAA	1362
Db	1385	GCAGAAATGATCAAGAGCTGGAATGACAGCGTTTAAAGACATGTCAGAGTGGCTTTGGAAA	1444
OY	1363	GGTGGTGGGGATCCAGTTCAAGAGAGGAGGGGTATGTTTGTCTCCACAGTCTGGGCAAGGA	1422
Db	1445	GGTGGTGGGGATCCAGTTCAAGAGAGGAGGGGTATGTTTGTCTCCACAGTCTGGGCAAGGA	1504
OY	1423	GTGCTATGCGGAATTCCTCTCATAGGAGAAAAGCTCCCAACATTTTCTTTGATGATGGTT	1482
Db	1505	GTGCTATGCGGAATTCCTCTCATAGGAGAAAAGCTCCCAACATTTTCTTTGATGATGGTT	1564
OY	1483	TTAAAGTCTCAGGCTTCTATAATAGAAAACAGAGGCTTCTTGACGCTCCTTGTTGGCGTG	1542
Db	1565	TTAAAGTCTCAGGCTTCTATAATAGAAAACAGAGGCTTCTTGACGCTCCTTGTTGGCGTG	1624
OY	1543	ATGTGCTCGAAATGATGTAAGTTCACGAGAAAGCATTTTTTTTTTCTTTGAACCTTAAGAT	1602

Db	1635	ATGTCCTCGAATATGATGTCACAGAAAGCAATTTTTCCTTGAACCTTAAAGCT	1684
Qy	1603	TCTATTATTAAGAGACAGATTCACATTTTATACATGAGATCTCTTGTGCGT	1662
Db	1685	TCTATTATTAAGAGACAGACATTCACATTTTATACATGAGATCTCTTGTGCGT	1744
Qy	1663	AATACCAGATTTAGCTGCATCCCTTTAAAGAAGTTTATATGTCCTGACTGTGGCTAAA	1722
Db	1745	AATACCAGATTTAGCTGCATCCCTTT - AAAGAAGTTTATGTCCTGACTGTGGCTAAA	1803
Qy	1723	TTATCTAATTTCCATGCTTTTGTAGATGATGAAGTATTTGTAGGCCACATATGGGA	1782
Db	1804	TTATCTAATTTCCATGCTTTTGTAGATGATGAAGTATTTGTAGGCCACATATGGGA	1863
Qy	1783	GTTCTAGATTTTGTAGTGAATGCGAGGAAAGGCCATCTCCATTCAGATGATTAAGTGAAAC	1842
Db	1864	GTTCTAGATTTTGTAGTGAATGCGAGGAAAGGCCATCTCCATTCAGATGATTAAGTGAAAC	1923
Qy	1843	AAACTAGTTCTCGGAATTTCTACAGAAAGGAGGAATCAGACTGAGAGAAAGCTGTGACATA	1902
Db	1924	AAACTAGTTCTCGGAATTTCTACAGAAAGGAGGAATCAGACTGAGAGAAAGCTGTGACATA	1983
Qy	1903	GGACTTGAAGCCCAAGACTTTGAAATTTGAGAGTGCATGCTGAGTTATATCACT	1962
Db	1984	GGACTTGAAGCCCAAGACTTTGAAATTTGAGAGTGCATGCTGAGTTATATCACT	2043
Qy	1963	GCTGCTTCTTCTATGAGTTACAAATCTATATTTTATTGAACTTTAAATTAAGAAAAAT	2022
Db	2044	GCTGCTTCTTCTATGAGTTACAAATCTATATTTTATTGAACTTTAAATTAAGAAAAAT	2103
Qy	2023	TTACAGAAAAAATAAAAAA 2041	
Db	2104	TTAAAAAATAAAAAAATAAAAAA 2122	
RESULT 3			
US-10-097-340-152			
Sequence 152, Application US/10097340			
GENERAL INFORMATION:			
APPLICANT: John MONAHAN			
APPLICANT: Manjula GANNANVARAPU			
APPLICANT: Sebastian HOERSCH			
APPLICANT: Shubhangi KAMATKAR			
APPLICANT: Steve G. KOVATS			
APPLICANT: Rachel E. MEYERS			
APPLICANT: Michael MORRISSEY			
APPLICANT: Peter OLANDT			
APPLICANT: Aml SEN			
APPLICANT: Peter VEIBY			
APPLICANT: Gordon B. MILLS			
APPLICANT: Robert C. BAST, Jr.			
APPLICANT: Karen LU			
APPLICANT: Rosemarie SCHMANDT			
APPLICANT: Xumei ZHAO			
APPLICANT: Karen GLAT			
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification			
FILE REFERENCE: Mri-030			
CURRENT APPLICATION NUMBER: US/10/097,340			
CURRENT FILING DATE: 2002-03-14			
PRIOR APPLICATION NUMBER: 60/276,025			
PRIOR FILING DATE: 2001-03-14			
PRIOR APPLICATION NUMBER: 60/325,149			
PRIOR FILING DATE: 2001-09-26			
PRIOR APPLICATION NUMBER: 60/276,026			
PRIOR FILING DATE: 2001-03-14			
PRIOR APPLICATION NUMBER: 60/324,967			
PRIOR FILING DATE: 2001/09/26			
PRIOR APPLICATION NUMBER: 60/311,732			
PRIOR FILING DATE: 2001-08-10			
PRIOR APPLICATION NUMBER: 60/325,102			
PRIOR FILING DATE: 2001-09-26			

QY	1021	TCCTGAATTCGTCAAGTTCTATCCAGGGGCCACAAGCTACAGCACTCCAAATAAAGT	1080
Db	1021	TCCTGAATTCGTCAAGTTCTATCCAGGGGCCACAAGCTACAGCACTCCAAATAAAGT	1080
QY	1081	GCCATACATGAAGCTGAGAACAAGAGATTATCCCTGTGTGACCTCTGTGATCGAAT	1140
Db	1081	GCCATACATGAAGCTGAGAACAAGAGATTATCCACCTGTGTGACCTCTGTGATCGAAT	1140
QY	1141	CATCATTTGGGGATTCGGGAATGGGCACGCCACATTAATTCCAATCCCACTTGAACCACT	1200
Db	1141	CATCATTTGGGGATTCGGGAATGGGCACGCCACATTAATTCCAATCCCACTTGAACCACT	1200
QY	1201	GAGAAAGAGAGAAATTGGACACACATCTCTCAACACCATAGCAATCCACAGTGTTC	1260
Db	1201	GAGAAAGAGAGAAAGTTGGACACACATCTCTCAACACCATAGCAATCCACAGTGTTC	1260
QY	1261	CCGACACTATACAAAGAACCTTAAGGGAAGGATCCCGAGGGCAGAAATGATCAAGACT	1320
Db	1261	CCGACACTATACAAAGAACCTTAAGGGAAGGATCCCGAGGGCAGAAATGATCAAGACT	1320
QY	1321	GAATGAGGGCTTAAGACACATGCTCAGTGGCCCTTGGAAAGGATGGGGGATTCAGTT	1380
Db	1321	GAATGAGGGCTTAAGACACATGCTCAGTGGCCCTTGGAAAGGATGGGGGATTCAGTT	1380
QY	1381	CAGGAGGAGGGGTATGTGTGTCTCCCACTCTGGCAAGAGAGTCTATGCGGAATTC	1440
Db	1381	CAGGAGGAGGGGTATGTGTGTCTCCCACTCTGGCAAGAGAGTCTATGCGGAATTC	1440
QY	1441	TGCATACAGAAAAGCTCCACCATTTCTTTTGAATGTGTTTAAGTCTCACGTTCTC	1500
Db	1441	TGCATACAGAAAAGCTCCACCATTTCTTTTGAATGTGTTTAAGTCTCACGTTCTC	1500
QY	1501	TATATATGAAGACGAGGTTGTGACGCTCTGTGAGTGTGAGTGTAAAGTCTCACGTTCTC	1560
Db	1501	TATATATGAAGACGAGGTTGTGACGCTCTGTGAGTGTGAGTGTAAAGTCTCACGTTCTC	1560
QY	1561	TAGTTCAAGAAACATTTTTTTTTCTTTGAACCTTAAGGTTCTATTATTAAGACAGC	1620
Db	1561	TAGTTCAAGAAACATTTTTTTTTCTTTGAACCTTAAGGTTCTATTATTAAGACAGC	1620
QY	1621	ACAAATTCACATTTTATATACATGAGGATCTCTTGTGTGAATACAGAGATTGACTGC	1680
Db	1621	ACAAATTCACATTTTATATACATGAGGATCTCTTGTGTGAATACAGAGATTGACTGC	1680
QY	1681	ATCCCTTTAAAGAAGTTTATGTCCCTACTCTGGCTAAATTTATCTAATTTCCAGATG	1740
Db	1681	ATCCCTTTAAAGAAGTTTATGTCCCTACTCTGGCTAAATTTATCTAATTTCCAGATG	1740
QY	1741	CTTTTGTAGATGACTGAATATTGTGTGAGCCACATATTGGAGTTCTAGATTTGAGTAA	1800
Db	1741	CTTTTGTAGATGACTGAATATTGTGTGAGCCACATATTGGAGTTCTAGATTTGAGTAA	1800
QY	1801	TGGGAGAAAGGGCCATCTCCATTGAGATGATTAAATGAACCAACTGTGTTCCGAAAT	1860
Db	1801	TGGGAGAAAGGGCCATCTCCATTGAGATGATTAAATGAACCAACTGTGTTCCGAAAT	1860
QY	1861	CTACAGAGAAAGGAGGAATCAGACTAGGAGAGCTGTGACATAGCACTTGAAGACCAAGA	1920
Db	1861	CTACAGAGAAAGGAGGAATCAGACTAGGAGAGCTGTGACATAGCACTTGAAGACCAAGA	1920
QY	1921	CTTTGAATTTGGAGCTGCTCATGTGTGATATTATATACAGTCTCTTCTATTTAGT	1980
Db	1921	CTTTGAATTTGGAGCTGCTCATGTGTGATATTATATACAGTCTCTTCTATTTAGT	1980
QY	1981	TACAATCTATATTATTATTTGAAGTTAAATAAGAAAAAATTTACAGAAAAA	2040
Db	1981	TACAATCTATATTATTATTTGAAGTTAAATAAGAAAAAATTTACAGAAAAA	2040
QY	2041	A 2041	
Db	2041	A 2041	

```

RESULT 2
PCT-US02-07826-152
; Sequence 152: Application PC/TUS0207826
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030PC
; CURRENT APPLICATION NUMBER: PCT-US02/07826
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 152
; LENGTH: 2129
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-07826-152

Query Match          94.6%; Score 1931; DB 1; Length 2129;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 2033; Conservative 1; Mismatches 6; Indels 79; Gaps 2

OY      1 CTGCCATAAGATGGCGTCCGTGGCGGCGTGCACAGACAGTTCTCTGTGGCGAGTGGCTCAG 60
DB      5 CTGCATTAAGATGGCGTCCGTGGCGGCTGCACAGACAGTTCTCTGTGGCGAGTGGCTCAG 64

OY      61 GGGCGCTGCAAGCGACCCCTACCTCTGTGTGTATCTCTCGGGGGCCAGGGGACCGGCAATC 120
DB      65 GGGCGCTGCAAGCGACCCCTACCTCTGTGTGTATCTCTCGGGGGCCAGGGGACCGGCAATC 124

OY      121 CACGCTGGCGTGTGACGTAGAGCCAGCGGCTCGCGGTGAGATGTCACGCGTACTCAT 180
DB      125 CACGCTGGCGTGTGACGTAGAGCCAGCGGCTCGCGGTGAGATGTCACGCGTACTCAT 184

OY      181 GCAGGCTATATGAAGGCTTAGACATATACCAACAAGTTTCTGCCAAGAGCAGAGAT 240
DB      185 GCAGGCTATATGAAGGCTTAGACATATACCAACAAGTTTCTGCCAAGAGCAGAGAT 244

OY      241 CTGCGGGGACCAATGATAGCTTTGTGTGATCCTCTTGACCAATTTACAGTGTGGA 300
DB      245 CTGCGGGGACCAATGATAGCTTTGTGTGATCCTCTTGACCAATTTACAGTGTGGA 304

OY      301 CTTCAGAAATAGAGCAACTGCTGTGATTTGAAGATATTTGGCCGAGACAAATTCAT 360
DB      305 CTTCAGAAATAGAGCAACTGCTGTGATTTGAAGATATTTGGCCGAGACAAATTCAT 364

OY      361 TGTGTGGGAGGAACCAATTATTTACATGAACTCTGCTCTGGAAAGTCTGTCAATAC 420
DB      365 TGTGTGGGAGGAACCAATTATTTACATGAACTCTGCTCTGGAAAGTCTGTCAATAC 424

OY      421 CAAGCCCCAGAGATGGGCACTGAGAAAGTATTTACCGCAAAAGTGAAGCTTGAAGA 480
DB      425 CAAGCCCCAGAGATGGGCACTGAGAAAGTATTTAGCCGAAAGTGAAGCTTGAAGA 484

OY      481 GGAATGCTTGTACTTACAAACGCCCTAAGCCAGGTGGAACCCAGAAATGGCTCCAAAGT 540
DB      485 GGAATGCTTGTACTTACAAACGCCCTAAGCCAGGTGGAACCCAGAAATGGCTCCAAAGT 544

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:26:48 ; Search time 6028.18 Seconds

(without alignments)
8512.654 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041
Sequence: 1 CTGCCATAGATGCGTCGCG.....TTTACACAGAAAAAAA 2041

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database : Pending Patents_NA_Main:*

```

1: /cgn2_6/ptodata/1/pna/US06.COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06.COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07.COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08.COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081.COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082.COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083.COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084.COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085.COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086.COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087.COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088.COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089.COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090.COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091.COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092.COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093.COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094.COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A.COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B.COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C.COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D.COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A.COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B.COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C.COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D.COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E.COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A.COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B.COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C.COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098A.COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US098B.COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098C.COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US099A.COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099B.COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099C.COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US099D.COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US100A.COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US100B.COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US101A.COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US101B.COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US102A.COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US102B.COMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2041	100.0	2041	19 US-09-513-151-3	Sequence 152, App
2	1931	94.6	2129	1 PCT-US02-07826-132	Sequence 152, App
3	1931	94.6	2129	39 US-10-097-340-152	Sequence 150, App
4	1929.4	94.5	2129	1 PCT-US02-07826-130	Sequence 150, App
5	1929.4	94.5	2129	39 US-10-097-340-150	Sequence 150, App
6	1923.6	94.2	2108	40 US-10-133-013-165	Sequence 165, App
7	1909	92.8	2109	76 US-60-324-185-33690	Sequence 33690, A
8	1894.6	92.8	2101	61 US-60-172-365-26034	Sequence 26034, A
9	1646	80.6	1858	18 US-09-471-275-2814	Sequence 2814, A
10	1556.8	76.3	1749	1 PCT-US02-18947-1692	Sequence 1692, App
11	1556.8	76.3	1749	41 US-10-172-118-1692	Sequence 874, App
12	1086.8	53.3	1183	29 US-09-757-028-874	Sequence 874, App
13	1086.8	53.3	1183	42 US-10-722-911-874	Sequence 874, App
14	1065.6	52.2	1223	28 US-09-705-256A-6006	Sequence 6006, App
15	1065.6	52.2	1223	60 US-60-164-285-6006	Sequence 6006, App
16	1051.2	51.5	1203	80 US-60-360-207-2119	Sequence 2119, App
17	1023.8	50.2	1041	16 US-09-205-070-8140	Sequence 8140, App
18	1023.8	50.2	1041	33 US-09-340-623-8140	Sequence 8140, App
19	1023.8	50.2	1041	33 US-09-898-888A-8140	Sequence 8140, App
20	1023.8	50.2	1041	33 US-09-898-888A-8140	Sequence 8140, App
21	1012.8	49.6	1039	17 US-09-359-922-3465	Sequence 3465, App

XX Claim 25; SEQ ID No 1507; 487pp; English.
PS
v

CC The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

50 Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;

Query Match	7.18; Score 145; DB 22; Length 457;
Best local similarity	100.0%; Posed No. 1 As-20.

Best Local Similarity 100.0%; Pred. No. 1.4e-29;

Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 AGGTCTATGAAGGCTAGACATCATCACCACAAGTTTCTGCCCAAGAGCAGAGAATCT 242

Db 450 AGGTCATGAAGCCTAGACATCATCAGCAACAAGTTTCTGCGCAAGAGCAGAGAACT 391

243 GCCGGCACCATGATCAGCTTTGTGGATCCTCTTGTGACCAATTACACAGTGTGACT 302

Db 390 GCGGCGCCACATGATCAGCTTGTGATCCTCTTGTGACCAATTACACAGTGTGACT 331

QY 303 TCAGAAATAGAGCAACTGCTCTGAT 327

Db 330 TCAGAAATAGAGCACTGCTCTGAT 306

Search completed: April 21, 2003, 13:37:09
Job time : 568.42 secs

```
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO: 1510; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;
XX
Query Match 7.1%; Score 145; DB 22; Length 457;
Best Local Similarity 100.0%; Pred. No. 1.4e-29;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 183 AGGCTATGAAGGCGCTAGACATCATCACCACACAGGTTTCTGCCAGAGCAGAGAATCT 242
DB 450 AGGCTATGAAGGCGCTAGACATCATCACCACACAGGTTTCTGCCAGAGCAGAGAATCT 391
QY 243 GCCGGCACCACATGATCAGCTTTGTGGATCCTCTTGTGACCAATTACACAGTGTGGACT 302
DB 390 GCCGGCACCACATGATCAGCTTTGTGGATCCTCTTGTGACCAATTACACAGTGTGGACT 331
QY 303 TCAGAAATAGAGCAACTGCTGTGAT 327
DB 330 TCAGAAATAGAGCAACTGCTGTGAT 306
XX
RESULT 14
AAK26966/C
ID AAK26966 standard; DNA; 457 BP.
XX
AC AAK26966;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 1523.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO: 1523; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;
XX
Query Match 7.1%; Score 145; DB 22; Length 457;
Best Local Similarity 100.0%; Pred. No. 1.4e-29;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO: 1523; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;
XX
Query Match 7.1%; Score 145; DB 22; Length 457;
Best Local Similarity 100.0%; Pred. No. 1.4e-29;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 183 AGGCTATGAAGGCGCTAGACATCATCACCACACAGGTTTCTGCCAGAGCAGAGAATCT 242
DB 450 AGGCTATGAAGGCGCTAGACATCATCACCACACAGGTTTCTGCCAGAGCAGAGAATCT 391
QY 243 GCCGGCACCACATGATCAGCTTTGTGGATCCTCTTGTGACCAATTACACAGTGTGGACT 302
DB 390 GCCGGCACCACATGATCAGCTTTGTGGATCCTCTTGTGACCAATTACACAGTGTGGACT 331
QY 303 TCAGAAATAGAGCAACTGCTGTGAT 327
DB 330 TCAGAAATAGAGCAACTGCTGTGAT 306
XX
RESULT 15
AA11574/C
ID AA11574 standard; DNA; 457 BP.
XX
AC AA11574;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #1507 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells.
```

PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-496933/54.
 DR
 XX
 PT New spatially-addressable set of single exon nucleic acid probes,
 PT useful for measuring gene expression in sample derived from human
 PT breast, comprises number of single exon nucleic acid probes
 XX
 PS Claim 1: SEQ ID NO 1534; 327bp + sequence listing; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and Bt 474 cells. The method involves contacting
 CC the probes with a collection of detectably labelled nucleic acids
 CC derived from mRNA of human breast, and then measuring the label
 CC bound to each probe of the microarray. The probes are useful for
 CC verifying the expression of regions of genomic DNA predicted to
 CC encode proteins. They are useful for gene discovery, and for
 CC determining predisposition and/or prognosing breast disease. Gene
 CC expression analysis is useful for assessing the toxicity of chemical
 CC agents on cells. The microarray of this invention presents a far greater
 CC diversity of probes for measuring gene expression, with far less bias
 CC than expressed sequence tag microarrays. The method is suitable for
 CC rapid production of functional information from genomic sequence. The
 CC present sequence is a single exon nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other:
 XX
 Query Match 7.1%; Score 145; DB 22; Length 457;
 Best Local Similarity 100.0%; Pred. No. 1.4e-29;
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 183 AGGCTATGAAGGCGCTAGACATCATCAACAAGGTTTGTGCCAAGAGCAGAAATCT 242
 DB 450 AGGCTATGAAGGCGCTAGACATCATCAACAAGGTTTGTGCCAAGAGCAGAAATCT 391
 OY 243 GCCGGCACACATGATGACGTTTGTGATCCTCTTGTGACCAATTAACAGTGTGACT 302
 DB 390 GCCGGCACACATGATGACGTTTGTGATCCTCTTGTGACCAATTAACAGTGTGACT 331
 OY 303 TCAGAAATAGAGCAACTGCTCTGAT 327
 DB 330 TCAGAAATAGAGCAACTGCTCTGAT 306
 RESULT 12
 ABA53262/c
 ID ABA53262 standard; DNA; 457 BP.
 XX
 AC ABA53262;
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Human foetal liver single exon nucleic acid probe #1567.
 XX
 KM Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 KW
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX

PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632386.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483447/52.
 DR
 XX
 PT human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver
 XX
 PS Claim 1: SEQ ID NO 1567; 639bp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other:
 XX
 Query Match 7.1%; Score 145; DB 22; Length 457;
 Best Local Similarity 100.0%; Pred. No. 1.4e-29;
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 183 AGGCTATGAAGGCGCTAGACATCATCAACAAGGTTTGTGCCAAGAGCAGAAATCT 242
 DB 450 AGGCTATGAAGGCGCTAGACATCATCAACAAGGTTTGTGCCAAGAGCAGAAATCT 391
 OY 243 GCCGGCACACATGATGACGTTTGTGATCCTCTTGTGACCAATTAACAGTGTGACT 302
 DB 390 GCCGGCACACATGATGACGTTTGTGATCCTCTTGTGACCAATTAACAGTGTGACT 331
 OY 303 TCAGAAATAGAGCAACTGCTCTGAT 327
 DB 330 TCAGAAATAGAGCAACTGCTCTGAT 306
 RESULT 13
 AAK01519/c
 ID AAK01519 standard; DNA; 457 BP.
 XX
 AC AAK01519;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe SEQ ID NO: 1510.
 XX
 KM Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 XX

KM pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US05171.
XX 17-FEB-2000; 2000US-183319P.
XX 16-MAR-2000; 2000US-189862P.
XX 25-MAY-2000; 2000US-207454P.
XX 09-JUN-2000; 2000US-211314P.
XX 18-JUL-2000; 2000US-219007P.
XX 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
XX Claim 1; Page 3210; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 545 BP; 179 A; 94 C; 91 G; 144 T; 37 other:
SQ
Query Match 8.0%; Score 163; DB 23; Length 545;
Best Local Similarity 79.1%; Pred. No. 1.7e-34;
Matches 193; Conservative 0; Mismatches 49; Indels 2; Gaps 1;
OY 1800 ATGCAGAGAAAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAACAGTTCGCGAAT 1859
DB 20 AAGCGGGNAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAACAGTTCGCGAAT 79
OY 1860 TCTACAGAGAGAGGAGATCAGACTGAGAGAGTGTGACATAGACTGAAGCAACAAAG 1919
DB 80 TTTTACACAGAGAGGAGATTAACCTGAGAGAGCTGTGACATAGACTGAAGCAACAAAG 139
OY 1920 ACTTGAATTTGGAGCTGCTCATGTGTGATTAATATCACTGCTCTTTCTATTGAG 1979
DB 140 ACTTGAATTTGGAGCTGCTCATGTGTGATTAATATCACTGCTCTTTCTATTGAG 199
OY 1980 TTAC-AAATCTATTTTATTTATTAAGTTTAATTAAGAAATTTTACAGAAAAAAA 2037
DB 200 TTACAAATTTTATTTTATTTATTAAGTTTATTAAGAAATTTTACAGCAACATCAAC 259
OY 2038 AAAA 2041
DB 260 AAAA 263
RESULT 9

ABV49413
ID ABV49413 standard; CDNA; 582 BP.
XX
XX AC ABV49413;
XX 17-SEP-2002 (first entry)
XX
XX Human prostate expression marker CDNA 49404.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US05171.
XX 17-FEB-2000; 2000US-183319P.
XX 16-MAR-2000; 2000US-189862P.
XX 25-MAY-2000; 2000US-207454P.
XX 09-JUN-2000; 2000US-211314P.
XX 18-JUL-2000; 2000US-219007P.
XX 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
XX Claim 1; Page 9652; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 582 BP; 201 A; 114 C; 111 G; 156 T; 0 other:
SQ
Query Match 7.9%; Score 162; DB 23; Length 582;
Best Local Similarity 84.7%; Pred. No. 3.3e-34;
Matches 205; Conservative 0; Mismatches 35; Indels 2; Gaps 2;
OY 1800 ATGCAGAGAAAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAACAGTTCGCGAAT 1859
DB 61 AAGCGGGNAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAACAGTTCGCGAAT 120
OY 1860 TCTACAGAGAGAGGAGATCAGACTGAGAGAGTGTGACATAGACTGAAGCAACAAAG 1919
DB 121 TCTACACAGAGAGGAGATGAGACTGAGAGAGCTGTACATAGACTGAAGCAACAAAG 180
OY 1920 ACTTGAATTTGGAGCTGCTCATGTGTG-AGTATTAATGACATGCTCTTTCTATTGGA 1978
DB 181 ACTTGAATTTGGAGCTGCTCATGTGTGAGACTGATATCACTGCTCTTTCTATTGGA 240

expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. MI can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of MI, and the data is sufficient to convey the chemical structure and/or properties of the agent. MI can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

Sequence 355 BP; 112 A; 74 C; 54 G; 115 T; 0 other:

Query Match 16.8%; Score 342.6; DB 24; Length 355;
Best Local Similarity 98.9%; Pred. No. 4e-84;
Matches 345; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1684 CCTTAAAGAGTTTATGTCCTGACCTGCTAAATTTCTATTTCCAGATGCTT 1743
1744 TTGTAGATGACGATGATTTGTGAGCCCATATTGGAGTCTGATTTGAGTAAAG 1803
289 TTGTGATGACGATGATTTGTGAGCCCATATTGGAGTCTGATTTGAGTAAAG 230
1804 CAGGAAGGGCCATCTCATTGATGATGATTAAGTAAAGCAACCTAGTTCCGAATCTTA 1863
229 CAGGAAGGGCCATCTCATTGATGATGATTAAGTAAAGCAACCTAGTTCCGAATCTTA 170
1864 CAGGAAGGGAGGATCAGACTGAGAGAGCTGTGACATGAGCTTGAACCAAGACTT 1923
169 CAGGAAGGGAGGATCAGACTGAGAGAGCTGTGACATGAGCTTGAACCAAGACTT 110
1924 TGAATTTGCGAGCTGCTCATGCTGAGTTTATTCACAGCTGCTTTTATTTAGTTAC 1983
109 TGAATTTGCGAGCTGCTCATGCTGAGTTTATTCACAGCTGCTTTTATTTAGTTAC 50
1984 AAATCTATATTTTATTTGAAGTTTAAATTAAGAAAAAATTTTCAAGAAA 2032
49 AAATCTATATTTTATTTGAAGTTTAAATTAAGAAAAAATTTTCAAGAAA 1

RESULT 7

AAA00911 standard; cDNA; 300 BP.

AAA00911;
19-MAY-2000 (first entry)

Human colon cancer cell line polynucleotide sequence SEQ ID NO:902.

Human: colon cancer; tumour; diagnosis; gene expression product;
probe; detection; cancerous state; metastasis; identification;
breast cancer; oestrogen receptor-positive breast cancer; therapy;
oestrogen receptor-negative breast cancer; lung cancer; ss.

Homo sapiens.

MO9958675-A2.

18-NOV-1999.

13-MAY-1999; 99WO-US10602.

14-MAY-1998; 98US-0085426.

15-MAY-1998; 98US-0085537.

15-MAY-1998; 98US-0085696.

21-OCT-1998; 98US-0105234.

27-OCT-1998; 98US-0105877.

(CHIR) CHIRON CORP.
(HSE-) HSE INC.

Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
WPI; 2000-126369/11.

Polynucleotide library used to determine cancerous states of mammalian cells -
Claim 1; Page 399; 1097pp; English.

AAA00010 to AAA02716 represent polynucleotides isolated from cDNA libraries constructed from human colon cancer cell lines. The present invention also describes a method of detecting differentially expressed genes correlated with a cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample derived from a cell suspected of being cancerous, where detection of the differentially expressed gene product is correlated with a cancerous state of the cell from which the test sample was derived. The polynucleotide sequences can be used in a method for detecting differentially expressed genes correlated with a cancerous state of a mammalian cell. The polynucleotides can also be used as probes for detecting and mapping related genes. They can be used in diagnosis and prognosis of diseases and disorders (e.g. identification of pre-metastatic or metastatic cancerous states, stages of cancer, or responsiveness of cancer to therapy). This is particularly for breast cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-negative breast cancer, lung cancer, and colon cancer.

Sequence 300 BP; 83 A; 69 C; 72 G; 76 T; 0 other;

Query Match 11.1%; Score 226.8; DB 21; Length 300;
Best Local Similarity 95.1%; Pred. No. 3.1e-52;
Matches 234; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

919 TAACAGCTTCTAAAGAAAGACAGCTGTCCTGTCCTGCTAGCTAGAGT 978
55 TAAAGACGTTTGTGAGAGAGACCTGGTCCATTTGCCCCCTGCTATGAGAGT 114
979 ATCTGATGCTCGAAGTGGAGAGCTGTCTTGAACCTGCTTGAATCGTCAAG 1038
115 ATCTGATGCTCGAAGTGGAGAGCTGTCTTGAACCTGCTTGAATCGTCAAG 174
1039 TTTTCATCCAGGGCCCAAGCCCTACAGCCCATTAAGAAAGCATCAAGAGCTGA 1098
175 TTTTCATCCAGGGCCCAAGCCCTACAGCCCATTAAGAAAGCATCAAGAGCTGA 234
1099 GAACAAGAGAACTATACCTGCTGAGCTGTGATGCAATCATTCATGGAGATCGGA 1158
235 GAACAAGAGAACTATACCTGCTGAGCTGTGATGCAATCATTCATGGAGATCGGA 294
QY 1159 ATGGGC 1164
DB 295 ATGGGC 300

RESULT 8

ABV19646 standard; cDNA; 545 BP.

ABV19646;

13-SEP-2002 (first entry)

Human prostate expression marker cDNA 19637.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

CC blood-related disorders (e.g. haemophilia), reproductive disorders
 CC (e.g. infertility) and infectious disorders (e.g. Influenza). The
 CC polynucleotides of the invention can also be used in gene therapy.
 CC AA40785-AA51684 represent cDNA sequences encoding for the novel human
 CC enzyme polypeptides of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 717 BP, 184 A; 177 C; 187 G; 162 T; 7 other;

Query Match 32.4%; Score 661.8; DB 22; Length 717;

Best Local Similarity 98.4%; Pred. No. 5.6e-172;

Matches 685; Conservative 2; Mismatches 7; Indels 2; Gaps 2;

OY 1 CTGCCATAAGATGCGCTCCGCGGCGCTGACAGACAGTCTCTGTGGGCGAGTGGCTCAG 60
 DB 1 CTGCCATAAGATGCGCTCCGCGGCGCTGACAGACAGTCTCTGTGGGCGAGTGGCTCAG 60
 OY 61 GGGCGCTGCAAGCGACCTACCTCTGTAGTGTCTCGGGGCGCAGCGGCGCAAAATC 120
 DB 61 GGGCGCTGCAAGCGACCTACCTCTGTAGTGTCTCGGGGCGCAGCGGCGCAAAATC 120
 OY 121 CACGCTGCGCTGCAAGCGCTGACAGCGGCTGAGATGCTGACGCTGACTCCAT 180
 DB 121 CACGCTGCGCTGCAAGCGCTGACAGCGGCTGAGATGCTGACGCTGACTCCAT 180
 OY 181 GCAGGTCTATGAGAGCGCTGACATCATCACAACAAGTTTCTGCCCAAGCAGAGAT 240
 DB 181 GCAGGTCTATGAGAGCGCTGACATCATCACAACAAGTTTCTGCCCAAGCAGAGAT 240
 OY 241 CCGCGGCGACCATGATGATGCTTGTGATGCTCTGTGACCAATTAACAGTGTGTA 300
 DB 241 CCGCGGCGACCATGATGATGCTTGTGATGCTCTGTGACCAATTAACAGTGTGTA 300
 OY 301 CTTGAGAAATAGAGCAACTGCTCTGATGAGATATATTGGCCGAGCAAAATTCCTAT 360
 DB 301 CTTGAGAAATAGAGCAACTGCTCTGATGAGATATATTGGCCGAGCAAAATTCCTAT 360
 OY 361 TTTTGTGGAGAGACCAATATTATTACATTGATCTGCTCTGGAAGTTCTGTCAATAC 420
 DB 361 TTTTGTGGAGAGACCAATATTATTACATTGATCTGCTCTGGAAGTTCTGTCAATAC 420
 OY 421 CAAGCGCGAGAGATGGGACCTGAGAAAGTGAATGACCAAAAGTGGAGCTTGAAGA 480
 DB 421 CAAGCGCGAGAGATGGGACCTGAGAAAGTGAATGACCAAAAGTGGAGCTTGAAGA 480
 OY 481 GGATGCTTGTACTTCAACAAAGCCTAAGCCAGGTGACCCAGAAATGGCTGCCAAGCT 540
 DB 481 GGATGCTTGTACTTCAACAAAGCCTAAGCCAGGTGACCCAGAAATGGCTGCCAAGCT 540
 OY 541 GCATCCACTGCAACAAAGCAAGTGGCGAGCTTGAAGTTTGAAGAAACAGAGAT 600
 DB 541 GCATCCACTGCAACAAAGCAAGTGGCGAGCTTGAAGTTTGAAGAAACAGAGAT 600
 OY 601 CTCTCATAGTGAATTTCTCATGTCACATATAACCGAAGAAAGTGTGTCCTCCNTTGA 660
 DB 601 CTCTCATAGTGAATTTCTCATGTCACATATAACCGAAGAAAGTGTGTCCTCCNTTGA 660
 OY 659 GGTCTCTGAGATTTCTTAACCCCTTGACCTTTGG 694
 DB 661 GGTCTCTGAGATTTCTTAACCCCTTGACCTTTTG 696

RESULT 6

ABL68710/c
 ID ABL68710 standard; DNA; 355 BP.

XX ABL68710;

DT 15-MAY-2002 (first entry)

XX Kidney cancer related gene sequence SEQ ID NO:7047.

XX Human: cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumor; carcinos;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 OS Homo sapiens.
 XX WO200194629-A2.
 PN 13-DEC-2001.
 PD 30-MAY-2001; 2001WO-US10838.
 PF 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 (AVAL-) AVALON PHARM.
 XX Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX MPI: 2002-188264/24.
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set
 XX
 PS Claim 1: SEQ ID 7047; 44pp; English.
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in

PR	17-NOV-2000	2000US-249209P
PR	17-NOV-2000	2000US-249210P
PR	17-NOV-2000	2000US-249211P
PR	17-NOV-2000	2000US-249212P
PR	17-NOV-2000	2000US-249213P
PR	17-NOV-2000	2000US-249214P
PR	17-NOV-2000	2000US-249215P
PR	17-NOV-2000	2000US-249216P
PR	17-NOV-2000	2000US-249217P
PR	17-NOV-2000	2000US-249218P
PR	17-NOV-2000	2000US-249244P
PR	17-NOV-2000	2000US-249245P
PR	17-NOV-2000	2000US-249264P
PR	17-NOV-2000	2000US-249265P
PR	17-NOV-2000	2000US-249297P
PR	17-NOV-2000	2000US-249299P
PR	17-NOV-2000	2000US-249300P
PR	01-DEC-2000	2000US-250160P
PR	01-DEC-2000	2000US-250391P
PR	05-DEC-2000	2000US-251030P
PR	05-DEC-2000	2000US-251988P
PR	05-DEC-2000	2000US-256719P
PR	06-DEC-2000	2000US-251479P
PR	08-DEC-2000	2000US-251865P
PR	08-DEC-2000	2000US-251868P
PR	08-DEC-2000	2000US-251869P
PR	08-DEC-2000	2000US-251889P
PR	08-DEC-2000	2000US-251900P
PR	11-DEC-2000	2000US-254097P
PR	05-JAN-2001	2001US-259678P

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-476161/51

P-PSDB; ABB10476.

PT	Isolated nucleic acid molecule encoding an inflammation-associated polypeptide is used in preventing, treating or ameliorating a medical condition -
----	--

Claim 1; SEQ ID NO: 364; 859pp + Sequence Listing; English..

CC The present invention provides human cDNAs, proteins and related genomic
CC DNAs. These can be used in the treatment of neural, immune system,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders and inflammation. The present sequence
CC is a cDNA of the invention.

50 Sequence 717 BP; 184 A; 177 C; 187 G; 162 T; 7 other;

Query Match	Score	DB	Length
32.48	661.8	22	717

Matches 685; Conservative 2; Mismatches 7; Indels 2; Gaps 2;

0Y 1 CTGCCATAGATGGCGTCCGTGGCGGCTGCACGAGCAGTTCCTGTGGGCAGTGGGCTCAG 60

Db 1 CTGCCATAAGATGGCGTCCGTGGCGGCTGCACGAGCAGTTCGCCGTGGGCAGTGGGCTCAG 60

61 GGGCCTGCACGGACCTACTTGTAGTATTCTCGGGCCACGGGCACCGCAATC 12

Db 61 GGGCCTGCACGGACCCTACTCTTGTAGTATCTCGGGGCCACGGGCACCGCAATC 12

QY 121 CACGCTGGCGTTGCAGCTAGGCCAGCGGCTCGGGGTGAGATCGTCAGCGCTGACTCCAT 18

Db 121 CACGCTGGCGTTGCAGCTAGGCCAGCGGCTCGGGGTGAGATCGTCAAGCGCTGACTCCAT 18

181 GCAGGCTATGAAGGCTAGACATCATCACCAACAAGTTTCTGCCCAAGAGCAGAGAAT 24

Db 181 GCAGGCTATGAAGGCTAGACATCATCACCACAAGGTTTCTGCCCAAGAGCAGAGAAT 24

241 CTGCCGGCACCACATGATCAGCTTTGTGGATCCTCTTGTGACCAATTACACAGTGTGA 30

Db	241	CTGCCGGCACCCACATGATCATGAGCTTTTGGATCCTCTTGTGACCAATATACAGAGTGGTGA	300
QY	301	CTTCAGAAATTAGACACACTGCTCTGATTGAGATATATTTGCCCGAGACAAATTCCTAT	360
Db	301	CTTCAGAAATTAGACACACTGCTCTGATTGAGATATATTTGCCCGAGACAAATTCCTAT	360
QY	361	TGTTGGGGAGGAGACCAATTATTACATTGAAATCTGCTCTGGAAGTTTGTGCAATAC	420
Db	361	TGTTGGGGAGGAGACCAATTATTACATTGAAATCTGCTCTGGAAGTTTGTGCAATAC	420
QY	421	CAAGCCCCGAGAGATGGGACCTGAGAAAGTATGATGACCGAAAGTGGAGCTTGAAAAAGA	480
Db	421	CAAGCCCCGAGAGATGGGACCTGAGAAAGTATGATGACCGAAAGTGGAGCTTGAAAAAGA	480
QY	481	GGATGGTCTTGTACTTCACAAAGCGCTTAAGCCAGTGGAGACCCGAAATGGCTGCCAAGCT	540
Db	481	GGATGGTCTTGTACTTCACAAAGCGCTTAAGCCAGTGGAGACCCGAAATGGCTGCCAAGCT	540
QY	541	GCATCCACATGACAAACGCCAAGGTGCCAGAGCTTGCAGATTTTGAAGAAACAGAGAT	600
Db	541	GCATCCACATGACAAACGCCAAGGTGCCAGAGCTTGCAGATTTTGAAGAAACAGAGAT	600
QY	601	CTCTCATAGTGAATTTCTTCATCGTCAACATA-CGGAAGAAGGTGGTGGTCCCC-TTGA	658
Db	601	CTCTCATAGTGAATTTCTTCATCGTCAACATACCGGAAGAAGGTGGTGGTCCCCNTTGA	660
QY	659	GGTCTCTGAGAGTCTCTTAACCCCTGGCATCTTTGG	694
Db	661	GGTCTCTGAGAGTCTCTTAACCCCTGGCATCTTTGG	696

RESULT 5
AAS1368

ID AAS41268 standard; cDNA; 717 BP.

AC AAS41268

DT 17-DEC-2001 (first entry)

CDNA encoding novel human enzyme polypeptide #484

KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;

KW autoimmune disorder; neurological disorder; metabolic disorder;

blood-related disorder; infectious disorder; gene therapy; cytostatic;

XX

XX
XX
XXXXXX

2
2
2
2
2
2
2

XX
E
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
8

XX	31 - JAN - 2000 - 2000hrs - 017000Z
DD	

PR 04-FEB-2000; 2000US-0180628.
PR 04-FEB-2000; 2000US-0184664

PR 02-MAR-2000; 2000US-0186350.
PR 15-MAR-2000; 2000US-0186374

PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0190133

PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0206457

PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.
PB 07-JUL-2000; 2000US-0216990

PR 11-JUL-2000; 2000US-0217487.
PB 11-JUL-2000; 2000US-0217486

PR 14-JUL-2000; 2000US-0218290.

QY 481 GGATGCTCTGCTACTGCACAAAGCCCTAAGCCAGGTGGACCCAGAAATGCTGCCAAGCT 540
Db 481 GGATGCTCTGCTACTGCACAAAGCCCTAAGCCAGGTGGACCCAGAAATGCTGCCAAGCT 540
QY 541 GCATCCACATGACAAAGCAAGAGTGGCCAGAGCTTGCAGATTTTGAAGAACAAGAAAT 600
Db 541 GCATCCACATGACAAAGCAAGAGTGGCCAGAGCTTGCAGATTTTGAAGAACAAGAAAT 600
QY 601 CTCCTCAGTGAATTTCTCCATCTCTCAACATA-CGGAAGAAGGTGGTGTCCCTTGG 658
Db 601 CTCCTCAGTGAATTTCTCCATCTCTCAACATA-CGGAAGAAGGTGGTGTCCCTTGG 660
QY 659 GGCTCTGGAAGTCTCTCAACCTTGCATCTTGG 694
Db 661 GGCTCTGGAAGTCTCTCAACCTTGCATCTTGG 696

RESULT 4
ABA06698
ID ABA06698 standard; cDNA; 717 BP.
XX ABA06698;
AC ABA06698;
XX 10-JAN-2002 (first entry)
DT 10-JAN-2002 (first entry)
XX
DE Human cDNA SEQ ID NO: 364.
XX
KW Human; gene therapy; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW proliferative disorder; inflammation; ss.
XX
OS Homo sapiens.
XX
PN W0200154474-A2.
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01349.
XX
PR 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 24-FEB-2000; 2000US-184664P.
PR 02-MAR-2000; 2000US-186350P.
PR 16-MAR-2000; 2000US-189874P.
PR 17-MAR-2000; 2000US-190076P.
PR 18-APR-2000; 2000US-198123P.
PR 19-MAY-2000; 2000US-205515P.
PR 07-JUN-2000; 2000US-209467P.
PR 28-JUN-2000; 2000US-214886P.
PR 30-JUN-2000; 2000US-215135P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225213P.
PR 14-AUG-2000; 2000US-225214P.
PR 14-AUG-2000; 2000US-225266P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 14-AUG-2000; 2000US-225759P.
PR 18-AUG-2000; 2000US-226279P.
PR 22-AUG-2000; 2000US-226811P.
PR 22-AUG-2000; 2000US-226868P.

PR 22-AUG-2000; 2000US-227183P.
PR 23-AUG-2000; 2000US-227009P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 06-SEP-2000; 2000US-230437P.
PR 06-SEP-2000; 2000US-230438P.
PR 08-SEP-2000; 2000US-231242P.
PR 08-SEP-2000; 2000US-231243P.
PR 08-SEP-2000; 2000US-231244P.
PR 08-SEP-2000; 2000US-231413P.
PR 08-SEP-2000; 2000US-231414P.
PR 08-SEP-2000; 2000US-232080P.
PR 08-SEP-2000; 2000US-232081P.
PR 12-SEP-2000; 2000US-231968P.
PR 14-SEP-2000; 2000US-233397P.
PR 14-SEP-2000; 2000US-233398P.
PR 14-SEP-2000; 2000US-233399P.
PR 14-SEP-2000; 2000US-233400P.
PR 14-SEP-2000; 2000US-233401P.
PR 14-SEP-2000; 2000US-233063P.
PR 14-SEP-2000; 2000US-233064P.
PR 14-SEP-2000; 2000US-233065P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 25-SEP-2000; 2000US-234998P.
PR 26-SEP-2000; 2000US-235484P.
PR 27-SEP-2000; 2000US-235834P.
PR 27-SEP-2000; 2000US-235836P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-238935P.
PR 13-OCT-2000; 2000US-238937P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241221P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241786P.
PR 20-OCT-2000; 2000US-241787P.
PR 20-OCT-2000; 2000US-241808P.
PR 20-OCT-2000; 2000US-241809P.
PR 20-OCT-2000; 2000US-241826P.
PR 01-NOV-2000; 2000US-244617P.
PR 08-NOV-2000; 2000US-246474P.
PR 08-NOV-2000; 2000US-246475P.
PR 08-NOV-2000; 2000US-246476P.
PR 08-NOV-2000; 2000US-246477P.
PR 08-NOV-2000; 2000US-246478P.
PR 08-NOV-2000; 2000US-246523P.
PR 08-NOV-2000; 2000US-246524P.
PR 08-NOV-2000; 2000US-246525P.
PR 08-NOV-2000; 2000US-246526P.
PR 08-NOV-2000; 2000US-246527P.
PR 08-NOV-2000; 2000US-246528P.
PR 08-NOV-2000; 2000US-246532P.
PR 08-NOV-2000; 2000US-246609P.
PR 08-NOV-2000; 2000US-246610P.
PR 08-NOV-2000; 2000US-246611P.
PR 17-NOV-2000; 2000US-249207P.
PR 17-NOV-2000; 2000US-249208P.

PR 08-SEP-2000; 2000US-231414P.
 PR 08-SEP-2000; 2000US-232080P.
 PR 08-SEP-2000; 2000US-232081P.
 PR 12-SEP-2000; 2000US-231968P.
 PR 14-SEP-2000; 2000US-232397P.
 PR 14-SEP-2000; 2000US-232398P.
 PR 14-SEP-2000; 2000US-232399P.
 PR 14-SEP-2000; 2000US-232400P.
 PR 14-SEP-2000; 2000US-232401P.
 PR 14-SEP-2000; 2000US-233063P.
 PR 14-SEP-2000; 2000US-233064P.
 PR 14-SEP-2000; 2000US-233065P.
 PR 21-SEP-2000; 2000US-234223P.
 PR 21-SEP-2000; 2000US-234224P.
 PR 25-SEP-2000; 2000US-234997P.
 PR 25-SEP-2000; 2000US-234998P.
 PR 25-SEP-2000; 2000US-235484P.
 PR 27-SEP-2000; 2000US-235834P.
 PR 27-SEP-2000; 2000US-235836P.
 PR 29-SEP-2000; 2000US-236327P.
 PR 29-SEP-2000; 2000US-236367P.
 PR 29-SEP-2000; 2000US-236368P.
 PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 02-OCT-2000; 2000US-237040P.
 PR 13-OCT-2000; 2000US-239935P.
 PR 13-OCT-2000; 2000US-239937P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241221P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241786P.
 PR 20-OCT-2000; 2000US-241787P.
 PR 20-OCT-2000; 2000US-241808P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 20-OCT-2000; 2000US-241826P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 08-NOV-2000; 2000US-246474P.
 PR 08-NOV-2000; 2000US-246475P.
 PR 08-NOV-2000; 2000US-246476P.
 PR 08-NOV-2000; 2000US-246477P.
 PR 08-NOV-2000; 2000US-246523P.
 PR 08-NOV-2000; 2000US-246524P.
 PR 08-NOV-2000; 2000US-246525P.
 PR 08-NOV-2000; 2000US-246526P.
 PR 08-NOV-2000; 2000US-246527P.
 PR 08-NOV-2000; 2000US-246528P.
 PR 08-NOV-2000; 2000US-246532P.
 PR 08-NOV-2000; 2000US-246609P.
 PR 08-NOV-2000; 2000US-246610P.
 PR 08-NOV-2000; 2000US-246611P.
 PR 08-NOV-2000; 2000US-246613P.
 PR 17-NOV-2000; 2000US-249207P.
 PR 17-NOV-2000; 2000US-249208P.
 PR 17-NOV-2000; 2000US-249209P.
 PR 17-NOV-2000; 2000US-249210P.
 PR 17-NOV-2000; 2000US-249211P.
 PR 17-NOV-2000; 2000US-249212P.
 PR 17-NOV-2000; 2000US-249213P.
 PR 17-NOV-2000; 2000US-249214P.
 PR 17-NOV-2000; 2000US-249215P.
 PR 17-NOV-2000; 2000US-249216P.
 PR 17-NOV-2000; 2000US-249217P.
 PR 17-NOV-2000; 2000US-249218P.
 PR 17-NOV-2000; 2000US-249244P.
 PR 17-NOV-2000; 2000US-249245P.
 PR 17-NOV-2000; 2000US-249246P.
 PR 17-NOV-2000; 2000US-249255P.
 PR 17-NOV-2000; 2000US-249297P.

PR 17-NOV-2000; 2000US-249299P.
 PR 17-NOV-2000; 2000US-249300P.
 PR 01-DEC-2000; 2000US-250160P.
 PR 01-DEC-2000; 2000US-250391P.
 PR 05-DEC-2000; 2000US-251030P.
 PR 05-DEC-2000; 2000US-251038P.
 PR 05-DEC-2000; 2000US-251088P.
 PR 05-DEC-2000; 2000US-251719P.
 PR 06-DEC-2000; 2000US-251479P.
 PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 PR 08-DEC-2000; 2000US-251989P.
 PR 08-DEC-2000; 2000US-251990P.
 PR 11-DEC-2000; 2000US-254097P.
 PR 05-JAN-2001; 2000US-259678P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX
 DR WPI: 2001-476161/51.
 DR P-PSDB: ABB10278.
 XX
 XX
 XX
 PT Isolated nucleic acid molecule encoding an inflammation-associated
 PT polypeptide is used in preventing, treating or ameliorating a medical
 PT condition
 PS
 PS Claim 1; SEQ ID NO: 166; 859pp + Sequence Listing; English.
 XX
 CC The present invention provides human cDNAs, proteins and related genomic
 CC DNAs. These can be used in the treatment of neural, immune system,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal and proliferative disorders and inflammation. The present sequence
 CC is a cDNA of the invention.
 XX
 SQ Sequence 717 BP; 184 A; 176 C; 187 G; 162 T; 8 other:
 Query Match 32.5%; Score 663; DB 22; Length 717;
 Best Local Similarity 98.4%; Pred. No. 2, 6e-172;
 Matches 685; Conservative 3; Mismatches 6; Indels 2; Gaps 2;
 QY 1 CTGCCATAGATGGCGTCGTCGGGCTGCACGAGCAAGTCTGTGGGCACTGGCTCAG 60
 DB 1 CTGCCATAGATGGCGTCGTCGGGCTGCACGAGCAAGTCTGTGGGCACTGGCTCAG 60
 QY 61 GGGCTTGCACGAGCACTTCTTGTAGTATTTCTGGGGCCACGGGCACTGGCAATC 120
 DB 61 GGGCTTGCACGAGCACTTCTTGTAGTATTTCTGGGGCCACGGGCACTGGCAATC 120
 QY 121 CAGCGTGGCGTTGAGCTAGGCGCAGGGGCTGGGGGAGATGTCAGGCGTGTACAT 180
 DB 121 CAGCGTGGCGTTGAGCTAGGCGCAGGGGCTGGGGGAGATGTCAGGCGTGTACAT 180
 QY 181 GCAGGTATGAAAGCGCTAGACATCATCACCACAAGGTTTCTGCCAAGAGCAGAGAT 240
 DB 181 GCAGGTATGAAAGCGCTAGACATCATCACCACAAGGTTTCTGCCAAGAGCAGAGAT 240
 QY 241 CTGCGGCGACCATGATGATCAGCTTGTGATCTCTTGTGCAATTAACAGTGTTGA 300
 DB 241 CTGCGGCGACCATGATGATCAGCTTGTGATCTCTTGTGCAATTAACAGTGTTGA 300
 QY 301 CTTGAGAAATGAGCACTGCTGATGTAAGATATATTTGCCGAGCAAAATTCCTAT 360
 DB 301 CTTGAGAAATGAGCACTGCTGATGTAAGATATATTTGCCGAGCAAAATTCCTAT 360
 QY 361 TGTGTGGAGAGCAATATTATTCATTGATCTCTCTGTAAGTTCTTGTCAATAC 420
 DB 361 TGTGTGGAGAGCAATATTATTCATTGATCTCTCTGTAAGTTCTTGTCAATAC 420
 QY 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTGATTTGACCCAAAGTGAAGCTTGAAGA 480
 DB 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTGATTTGACCCAAAGTGAAGCTTGAAGA 480

ABAO6500	standard; cDNA; 717 bp.
ABAO6500	
ABAO6500;	
10-JAN-2002	(first entry)
Human cDNA SEQ ID NO: 166.	
Human; gene therapy; neural disorder; immune system disorder;	
muscular disorder; reproductive disorder; gastrointestinal disorder	
pulmonary disorder; cardiovascular disorder; renal disorder;	
proliferative disorder; inflammation; ss.	
Homo sapiens.	
MO200154474-A2.	
02-AUG-2001.	
17-JAN-2001;	2001MO-US01349.
31-JAN-2000;	2000US-179065P.
04-FEB-2000;	2000US-180628P.
24-FEB-2000;	2000US-184664P.
02-MAR-2000;	2000US-186360P.
16-MAR-2000;	2000US-189874P.
17-MAR-2000;	2000US-190076P.
18-APR-2000;	2000US-198123P.
19-MAY-2000;	2000US-205515P.
07-JUN-2000;	2000US-209467P.
28-JUN-2000;	2000US-214886P.
30-JUN-2000;	2000US-215135P.
07-JUL-2000;	2000US-216647P.
07-JUL-2000;	2000US-216880P.
11-JUL-2000;	2000US-217487P.
11-JUL-2000;	2000US-217496P.
14-JUL-2000;	2000US-218290P.
26-JUL-2000;	2000US-220963P.
26-JUL-2000;	2000US-220964P.
14-AUG-2000;	2000US-224518P.
14-AUG-2000;	2000US-224519P.
14-AUG-2000;	2000US-225213P.
14-AUG-2000;	2000US-225214P.
14-AUG-2000;	2000US-225266P.
14-AUG-2000;	2000US-225267P.
14-AUG-2000;	2000US-225268P.
14-AUG-2000;	2000US-225270P.
14-AUG-2000;	2000US-225447P.
14-AUG-2000;	2000US-225757P.
14-AUG-2000;	2000US-225758P.
14-AUG-2000;	2000US-225759P.
18-AUG-2000;	2000US-226279P.
22-AUG-2000;	2000US-226681P.
22-AUG-2000;	2000US-226688P.
22-AUG-2000;	2000US-227182P.
23-AUG-2000;	2000US-227009P.
30-AUG-2000;	2000US-228924P.
01-SEP-2000;	2000US-229287P.
01-SEP-2000;	2000US-229343P.
01-SEP-2000;	2000US-229344P.
01-SEP-2000;	2000US-229345P.
05-SEP-2000;	2000US-229509P.
05-SEP-2000;	2000US-229513P.
06-SEP-2000;	2000US-230437P.
06-SEP-2000;	2000US-230438P.
08-SEP-2000;	2000US-231242P.
08-SEP-2000;	2000US-231243P.
08-SEP-2000;	2000US-231244P.
08-SEP-2000;	2000US-231413P.

Df	1801	TGCGAGAAAGGGCACCCTCCATTCATTGACATGATTAAAGCAACCAACTAGTCTCGGAATT	1860
OY	1861	CTACAGAGAAGAGAGGAATACAGCTGAGAGAGCTGTGCACATAGGACTTGAGAACCAAAGA	1920
Df	1861	CTACAGAGAAGAGAGGAATACAGCTGAGAGAGCTGTGCACATAGGACTTGAGAACCAAAGA	1920
OY	1921	CTTTGAAATTTGGAGCTGCTCATNGTGTGAGTATTAATACACCTGCTCTTTCATATGAGCT	1980
Df	1921	CTTTGAAATTTGGAGCTGCTCATNGTGTGAGTATTAATACACCTGCTCTTTCATATGAGCT	1980
OY	1981	TACAAATCTATATTTTATTTGAGTTAATAATAAGAAAAATTTACAAGAAAAAAA	2040
Df	1981	TACAAATCTATATTTTATTTGAGTTAATAATAAGAAAAATTTACAAGAAAAAAA	2040
OY	2041	A 2041	
Df	2041	A 2041	
 RESULT 2 ABN59708			
ID	ABN59708	standard; cDNA; 2130 BP.	
XX	XX		
AC	ABN59708;		
XX	XX		
DY	28-JUN-2002	(first entry)	
XX	XX		
DE	Novel human coding sequence SEQ ID NO: 119.		
XX	XX		
KW	Human; antihaemetic; vulnerary; antiinflammatory; immunomodulator;		
MW	antileptility; cerebroprotective; cyostatic; rheumatic; gene therapy;		
KM	neuroprotective; antiparkinsonian; protein therapy; EST;		
XX	expressed sequence tag; gene; ss.		
OS	Homo sapiens.		
PN	WO200222660-A2.		
PD	21-MAR-2002.		
PF	10-SEP-2001; 2001WO-US26015.		
PR	11-SEP-2000; 2000US-0659671.		
PA	(HYSE-) HYSEQ INC.		
PI	Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;		
DR	Xue AJ, Yang Y, Wehrman T, Dimaac RT;		
WI	2002-292408/33.		
P-PSDB	ABB97295.		
PT	An isolated polynucleotide for treating diseases associated with its		
XX	encoded polypeptide such as cancer and multiple sclerosis -		
PS	Claim 1; SEQ ID NO 119; 509bp; English.		
CC	The present invention provides the protein and coding sequences of 444		
CC	novel human proteins. These were isolated from expressed sequences tags		
CC	(ESTs). They can be used to stimulate cell growth, to regulate		
CC	hematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth		
CC	e.g. in burn treatment, to regulate the immune system e.g. to treat		
CC	multiple sclerosis, to regulate activin or inhibin e.g. to treat		
CC	fertility, to regulate haemostasis or thrombolysis e.g. to treat		
CC	stroke and cancer, to screen for drugs, to treat inflammatory conditions		
CC	e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.		
CC	Parkinson's disease. The present sequence is a coding sequence of the		
CC	invention.		
XX	Sequence 2130 BP; 620 A; 439 C; 522 G; 549 T; 0 other;		
XX	Query Match	94.1%; Score 1921.2; DB 24; Length 2130;	

	Best Local Similarity	95.9%	Pred. No.	0:	Mismatches	8:	Indels	79:	Gaps	2:
	Matches	2025:	Conservative	0:						
OY	8	AAGATGGCGTCCGTTGGGCGGTGCACAGACAGTTCCTGTGGGCAAGTGCGCTCAGGGGCGCTG	67							
Db	20	ACGAATTCCGTCGCTGGCGGCTGCACAGACAGTATCCCGTTGGCGCATGTGGGCTCAGGGGCGCTG	79							
OY	68	CACGGACCCTTAACCTCTTGTAAGTAGTTTCCGGGGGCGACGGGACCGGGCAATTCACACGCTG	127							
Db	80	CAACGACCCTTAACCTCTTGTAAGTATTCCTGGGGGCGACGGGACCGGGCAAAATTCACACGCTG	139							
OY	128	GGCTTGCAGACTAAGGCGCACGGGCTCGGGGTTGAGATCTGCACGCGCTGACTCCATGCAAGTTC	187							
Db	140	GGCTTGCAGACTAAGGCGCACGGGCTCGGGGTTGAGATCTGCACGCGCTGACTCCATGCAAGTTC	199							
OY	188	TATGAAGGCTTAGACATCATCACCAAGAAGTTTTGCCCCAAGACAGACAAGAAATCTGCCG	247							
Db	200	TATGAAGGCTTAGACATCATCACCAAGAAGTTTTGCCCCAAGACAGACAAGAAATCTGCCG	259							
OY	248	CACCATGATGACGTTTGTGGATCTCTTGTGACCAATTACACAGTGTGGACTTCAGA	307							
Db	260	CACCATGATGACGTTTGTGGATCTCTTGTGACCAATTACACAGTGTGGACTTCAGA	319							
OY	308	AATAAGCAACGCTCTGATTTGAATATATTTGGCCCGAGACAATAATCTATTTGTTGG	367							
Db	320	AATAAGCAACGCTCTGATTTGAATATATTTGGCCCGAGACAATAATCTATTTGTTGG	379							
OY	368	GGAGGAACCAATTTATTAATTGAATCTCGCTGTGGAAAATTCCTTGTCATACCAAGCCC	427							
Db	380	GGAGGAACCAATTTATTAATTGAATCTCGCTGTGGAAAATTCCTTGTCATACCAAGCCC	439							
OY	428	CAGGAGATGGGCACTGAGAAAGTGATGACCGAAAAAGTGGAGCTTGAAAAAGAGATGGT	487							
Db	440	CAGGAGATGGGCACTGAGAAAGTGATGACCGAAAAAGTGGAGCTTGAAAAAGAGATGGT	499							
OY	488	CTTGACTTTCACAAAGCGCTAAGCCAAGGTGGACCCAGAAATGGCTGCCAACGTCGATCCA	547							
Db	500	CTTGACTTTCACAAAGCGCTAAGCCAAGGTGGACCCAGAAATGGCTGCCAACGTCGATCCA	559							
OY	548	CATGACAAACGCAAGTGGCGAGAGCTTGCACAGTTTGTGAAGAAAACAGAAATCTCTCAT	607							
Db	560	CATGACAAACGCAAGTGGCGAGAGCTTGCACAGTTTGTGAAGAAAACAGAAATCTCTCAT	619							
OY	608	AGTGAATTTCTCCATGCTCACATACGAGAGAAGGTGTGTCCTCCCTGGAGTCTCTCG	667							
Db	620	AGTGAATTTCTCCATGCTCACATACGAGAGAAGGTGTGTCCTCCCTGGAGTCTCTCG	679							
OY	668	AAGTCTCTAACCCTTGCACTCTTGGGCTTAANGTGACACAGGATTTAGATGAGGCC	727							
Db	680	AAGTCTCTAACCCTTGCACTCTTGGGCTTAANGTGACACAGGATTTAGATGAGGCC	739							
OY	728	TTGGATTAAGAGGGTGGATGACATGCTGTGCTGTGGGCTCTTGGAGGAATTAAGAGATTTT	787							
Db	740	TTGGATTAAGAGGGTGGATGACATGCTGTGCTGTGGGCTCTTGGAGGAATTAAGAGATTTT	799							
OY	788	CACAGACGCTATATACAGAGAATGTTTCGAAAAATAGCCAGAGACTATACACATGGTATC	847							
Db	800	CACAGACGCTATATACAGAGAATGTTTCGAAAAATAGCCAGAGACTATACACATGGTATC	859							
OY	848	TTCCAAATCAATTTGGCTTCAAGGAATTTCAACGAGTACCTGATACCTGAGGGAATAATGCCACA	907							
Db	860	TTCCAAATCAATTTGGCTTCAAGGAATTTCAACGAGTACCTGATACCTGAGGGAATAATGCCACA	919							
OY	908	CTGGAGACTAGTAACACAGGTTCTTAAGAAAG-----	938							
Db	920	CTGGAGACTAGTAACACAGGTTCTTAAGAAAGATTTGAGGCTCTGAACAACAATAACTAAG	979							
OY	939	-----GACCTGGTCCC	949							
Db	980	AGATATGCCCGGAAACAACCGATGGGTTAAAAACCGTTTTTTTGACGACAGACTGGTCCC	1039							
OY	950	ATTGTCCCCCTGTCTATAGGCTTAGAGGTATCTGATGTCTCGAAGTGGAGAGACTCTGTT	1009							

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:03:18 ; Search time 557.42 Seconds

(without alignments)
8245.719 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041
Sequence: 1 CTCGCATAGATGCGTCCG.....TTTACAGAAAAA 2041

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2041	100.0	2041	20	AAK36073	Human homologue of
2	1921.2	94.1	2130	24	ABN59708	Human homologue of
3	663	32.5	717	22	ABN06500	Human CDNA SEQ ID
4	661.8	32.4	717	22	ABN06598	Human CDNA SEQ ID
5	661.8	32.4	717	22	AA541268	Human CDNA SEQ ID
6	342.6	16.8	355	24	ABN68710	CDNA encoding nove
7	226.8	11.1	300	21	AAA00911	Kidney cancer rela
8	163	8.0	545	23	ABN19646	Human colon cancer
9	162	7.9	582	23	ABV49413	Human prostate exp
						Human prostate exp

C	10	160.8	7.9	6399	23	ABN17248
C	11	145	7.1	457	22	ABN42839
C	12	145	7.1	457	22	ABN53262
C	13	145	7.1	457	22	AAK01519
C	14	145	7.1	457	22	AAK26966
C	15	145	7.1	457	22	AAI11574
C	16	145	7.1	457	22	AAI32862
C	17	138	6.8	138	22	ABN47966
C	18	138	6.8	138	22	ABN45854
C	19	138	6.8	138	22	AAK14265
C	20	138	6.8	138	22	AAK39928
C	21	138	6.8	138	22	AAI20787
C	22	138	6.8	138	22	AAI46014
C	23	75.8	3.7	6225	23	ABN10424
C	24	75.8	3.7	6225	23	ABN10424
C	25	75	3.7	903	22	AAH67098
C	26	75	3.7	1026	22	AAH67098
C	27	75	3.7	349980	22	AAH67098
C	28	73	3.6	1238	21	AAH68350
C	29	71.4	3.5	1243	21	AAH68350
C	30	71.4	3.5	1713	21	AAH68350
C	31	71.4	3.5	1713	21	AAH68350
C	32	69.6	3.4	700	21	AAH68350
C	33	65	3.2	582	24	ABN91871
C	34	65	3.2	990	22	AAH53033
C	35	65	3.2	3775	22	AAH54468
C	36	62.6	3.1	1374	21	AAH49935
C	37	61	3.0	1376	21	AAH49935
C	38	60	2.9	60	24	AAH49935
C	39	58.4	2.9	4411529	22	AAH49935
C	40	57.6	2.8	918	24	ABO69238
C	41	57.6	2.8	2944528	24	ABO69238
C	42	52	2.5	1092	24	ABO69242
C	43	52	2.5	495269	24	ABO69242
C	44	50.4	2.5	8519	20	AAH13043
C	45	49.8	2.4	989	18	AAV75111

ALIGNMENTS

RESULT 1
AAK36073 standard; DNA; 2041 BP.

XX AAK36073:

XX 15-JUL-1999 (first entry)

XX Human homologue of gro-1, referred to as hgro-1.

DE gro-1 operon; gro-1 gene; gop-1 gene; gop-2 gene; gop-2 gene;

XX hap-1 gene; Cancer; aging; longevity; tumour formation;

KW physiological clock; ss.

XX Homo sapiens.

OS WO9910482-A1.

XX 04-MAR-1999.

XX 20-AUG-1998; 98WO-CA00803.

XX 25-AUG-1997; 97CA-2210251.

XX (UYMC-) UNIV MCGILL.

XX Barnes T, Hekimi S, Lakowski B, Lemieux J;

PI WPI; 1999-190615/16.

DR Molecular identity of the gro-1 gene - useful for cancer diagnosis

XX and/or prognosis, and where compounds affecting encoded proteins are

PT

Db 196247 GAGGATCTTTTAAAGTGGTAATACCAAGATTCAGTGCAT-CCTTTAAAGAGCGTCTT 196305
OY 1703 GTCCCTGACTGCGGCTAAATTAATCTAATTCACAGATGCTTTTGTAGATGACTGAAGTAT 1762
||||| | ||||||| | | ||||||| ||||||| ||||||| |||||||
Db 196306 GTCCCTGGCGCTGGCTAAATAATAGCTCGTTCCAGATGCTTTGTAGATGACTGAAGTAT 196365
OY 1763 TTGTGAGCCACATATTGGGAGTCTTGTAGATTGTGATGTAATGGCAGGAAAGGCCATCTCCA 1822
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 196366 TGTAAAGGCGCAGTCAAGAGCTCTGAGCTCGAAGACGACGAAAGAGAGTAG-----TGCA 196419
OY 1823 TTGAGATGATTAACTGAACCAACTAGTCTCTCGAATTCTACAGAGAAAGAGGGAATCAG 1882
||||| ||||||| | | | | | | | | | | | | | | | | | | | | | |
Db 196420 GAGAGATGATTAAAGCAACTCTCCAGCTCTGTGAAT---CTACAGAAAGAGGGGCTCAG 196475
OY 1883 ACTGAGAAAGCTGTGACATAGAGACTTGAAGAC-CAAGACTTTGAATTTCGAGCTGCT 1941
||||| | |||||| | | ||||||| |||||| | | | |||||||
Db 196476 GCTGAGGTGTTGTGACCTGTGATCTTGAAGACTAAAGACACAGAGCTGCTGAGCTGCT 196535
OY 1942 CATGTG 1947
| | | |
Db 196536 COTGTG 196541

Search completed: April 21, 2003, 15:52:47
Job time : 7858.89 secs

[illegible]

*** 35 unordered pieces.

ACCESSION AC093939

VERSION AC093939.4 GI:21723477

KEYWORDS HTG; HTGS_PHASE1.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 239888)

AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alshrocks,S.L., Amaralunga,H.C., Are,J.R., Ayale,M., Banks,T., Barberia,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhan,C., Burck,C., Burckell,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwalte,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gablil,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homsl,F., Howard,S., Huber,J., Huylk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichteage,O., Lien,C., Liu,J., Liu,W., Lohseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheswari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Meizer,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Nguyen,N., Nickerson,E., Nwokenwo,S., Ogund,M., Okunolu,G., Otagunye,N., Oyiedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rivers,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Umami,K., Vasquez,L., Vera,Y., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wellington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G. and Gibbs,R.

TITLE Unpublished

JOURNAL 2 (bases 1 to 239888)

AUTHORS

Direct Submission
Submitted (12-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 239888)

AUTHORS

Direct Submission
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

JOURNAL

COMMENT

On Jul 10, 2002 this sequence version replaced gi:20335575.

Center: Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GEXO

Center clone name: CH230-172N14

Summary Statistics

Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 210507 bases at least Q40
Consensus quality: 213691 bases at least Q30
Consensus quality: 215787 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 35 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1	1220:	contig of 1220 bp in length
1221	1320:	gap of unknown length
1321	2370:	contig of 1050 bp in length
2371	2470:	gap of unknown length
2471	3511:	contig of 1041 bp in length
3512	3612:	gap of unknown length
3612	4645:	contig of 1034 bp in length
4646	4745:	gap of unknown length
4746	5993:	contig of 1248 bp in length
5994	6094:	gap of unknown length
6094	7118:	contig of 1025 bp in length
7119	7219:	gap of unknown length
7219	8699:	contig of 1381 bp in length
8699	8700:	gap of unknown length
8700	9998:	contig of 1299 bp in length
9999	10098:	gap of unknown length
10099	11207:	contig of 1109 bp in length
11208	11307:	gap of unknown length
11308	12832:	contig of 1525 bp in length
12833	12932:	gap of unknown length
12933	15681:	contig of 2749 bp in length
15682	15781:	gap of unknown length
15782	18002:	contig of 2221 bp in length
18003	18102:	gap of unknown length
18103	20292:	contig of 2190 bp in length
20293	20392:	gap of unknown length
20393	22189:	contig of 1797 bp in length
22190	22289:	gap of unknown length
22290	24818:	contig of 2529 bp in length
24819	24918:	gap of unknown length
24919	29103:	contig of 4185 bp in length
29104	29203:	gap of unknown length
29204	31883:	contig of 2680 bp in length
31884	31883:	gap of unknown length
31984	34715:	contig of 2732 bp in length
34716	34815:	gap of unknown length
34816	39740:	contig of 4925 bp in length
39741	39840:	gap of unknown length
39841	42402:	contig of 2562 bp in length
42403	42502:	gap of unknown length
42503	48347:	contig of 5845 bp in length
48348	48447:	gap of unknown length
48448	52407:	contig of 3960 bp in length
52408	52507:	gap of unknown length
52508	56832:	contig of 4325 bp in length
56833	56932:	gap of unknown length
56933	63189:	contig of 6257 bp in length
63190	63289:	gap of unknown length
63290	68850:	contig of 5561 bp in length
68851	68950:	gap of unknown length
68951	76141:	contig of 7191 bp in length
76142	76241:	gap of unknown length
76242	83247:	contig of 7006 bp in length
83248	83347:	gap of unknown length
83348	94069:	contig of 10722 bp in length

RESULT	13
AC093939	
LOCUS	AC093939
DEFINITION	Rattus norvegicus clone CH230-172N14. *** SEQUENCING IN PROGRESS
	23688 bp DNA linear HTG 11-JUL-2002

Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
 Mitter, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokwuo, S., Ogun, M., Okwona, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Pevens, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
 Scherer, S., Scott, G., Shen, H., Shooshari, N., Sisson, I.,
 Sodergren, E., Sotnik, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Swatek, A., Taber, P., Tamerisa, A., Tamerisa, R., Tang, H.,
 Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,
 Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.

Direct Submission
 2 (bases 1 to 181838)
 Morley, K.C.

Direct Submission
 Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 181838)
 Morley, K.C.

Direct Submission
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 11, 2002 this sequence version replaced gi:18139174.

----- Genome Center -----
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center Project name: GLTJ
 Center Clone name: CH230-64124
 Summary Statistics

Sequencing vector: M13;
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 115969 bases at least Q40
 Consensus quality: 122034 bases at least Q30
 Consensus quality: 127422 bases at least Q20

----- NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 62 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1073: contig of 1073 bp in length
 1074 1173: gap of unknown length
 1174 2367: contig of 1194 bp in length
 2368 2467: gap of unknown length
 2468 3755: contig of 1288 bp in length
 3756 3855: gap of unknown length
 3856 5138: contig of 1283 bp in length
 5139 5238: gap of unknown length
 5239 6254: contig of 1016 bp in length
 6255 7644: gap of unknown length
 7645 7744: contig of 1290 bp in length
 7745 9196: gap of unknown length
 9197 9296: gap of unknown length

9297 10423: contig of 1127 bp in length
 10424 10523: gap of unknown length
 10524 12093: contig of 1570 bp in length
 12094 12193: gap of unknown length
 12194 13606: contig of 1413 bp in length
 13607 15139: gap of unknown length
 15140 15239: gap of unknown length
 15240 16519: contig of 1280 bp in length
 16520 16620: gap of unknown length
 16621 18029: contig of 1410 bp in length
 18030 18129: gap of unknown length
 18130 19537: gap of unknown length
 19538 19637: gap of unknown length
 19638 20773: contig of 1136 bp in length
 20774 20873: gap of unknown length
 20874 21940: contig of 1067 bp in length
 21941 22040: gap of unknown length
 22041 23977: contig of 1937 bp in length
 23978 24077: gap of unknown length
 24078 25464: contig of 1387 bp in length
 25465 25564: gap of unknown length
 25565 26881: contig of 1317 bp in length
 26882 26981: gap of unknown length
 26982 28875: gap of unknown length
 28876 28975: gap of unknown length
 28976 30545: contig of 1570 bp in length
 30546 31672: contig of 1027 bp in length
 31673 31772: gap of unknown length
 31773 33195: contig of 1423 bp in length
 33196 33295: gap of unknown length
 33296 34682: contig of 1387 bp in length
 34683 34782: gap of unknown length
 34783 35965: contig of 1184 bp in length
 35966 36065: gap of unknown length
 36066 37755: gap of unknown length
 37756 37855: gap of unknown length
 37856 40161: contig of 2306 bp in length
 40162 40261: gap of unknown length
 40262 42177: contig of 1916 bp in length
 42178 42277: gap of unknown length
 42279 44344: contig of 2067 bp in length
 44345 44444: gap of unknown length
 44445 46472: contig of 2028 bp in length
 46473 46572: gap of unknown length
 46573 48054: contig of 1482 bp in length
 48055 48154: gap of unknown length
 48155 49390: contig of 1236 bp in length
 49391 49490: gap of unknown length
 49491 52986: contig of 3496 bp in length
 52987 53086: gap of unknown length
 53087 54647: contig of 1561 bp in length
 54648 54747: gap of unknown length
 54748 57112: contig of 2265 bp in length
 57113 57212: gap of unknown length
 57213 58488: contig of 1636 bp in length
 58489 58948: gap of unknown length
 58949 60817: contig of 1869 bp in length
 60818 60917: gap of unknown length
 60919 63443: contig of 2526 bp in length
 63444 63543: gap of unknown length
 63544 66294: contig of 2751 bp in length
 66295 66394: gap of unknown length
 66395 68140: contig of 1746 bp in length
 68141 68241: gap of unknown length
 68242 71198: contig of 2958 bp in length
 71199 71298: gap of unknown length
 71299 73990: contig of 2692 bp in length
 73991 74090: gap of unknown length
 74091 75872: contig of 1782 bp in length
 75873 75972: gap of unknown length
 75973 79069: contig of 3097 bp in length


```

QY 1211 AGAAGATTGGACTCAGATGCTGTCAACACCATAGAAAGTCAGATGTTTCCCGAGACTAT 1270
      |||
Db 93981 AGAAGATTGGACTCAGATGCTGTCAACACCATAGAAAGTCAGATGTTTCCCGAGACTAT 93922
QY 1271 AACAAAGAACCTAAAGGAGGATGCCCGAGGAGATGATCATCAGAGCTGAATGCAGC 1330
      |||
Db 93921 AACAAAGAACCTAAAGGAGGATGCCCGAGGAGATGATCATCAGAGCTGAATGCAGC 93862
QY 1331 GTTTAAGACATGTCACATGTCGCTTTGGAAGGTGGTGGGATCCAGTTCCAGAGGAG 1390
      |||
Db 93861 GTTTAAGACATGTCACATGTCGCTTTGGAAGGTGGTGGGATCCAGTTCCAGAGGAG 93802
QY 1391 GGCTATGTTTGTCTCCAGCTGCGGCAAGAGATGCTATCCGGAATTCCTGCATACAG 1450
      |||
Db 93801 GGCTATGTTTGTCTCCAGCTGCGGCAAGAGATGCTATCCGGAATTCCTGCATACAG 93742
QY 1451 AAAAGCTCCACCATTTTCTTTGATGCTTTTAAAGTCTCAGCTCTCTATATATGAA 1510
      |||
Db 93741 AAAAGCTCCACCATTTTCTTTGATGCTTTTAAAGTCTCAGCTCTCTATATATGAA 93682
QY 1511 ACAGACAGCTTGTCAAGCTCTTGTGTGCTGATGTGTGGAATGATGTACTTCAGA 1570
      |||
Db 93681 ACAGACAGCTTGTCAAGCTCTTGTGTGCTGATGTGTGGAATGATGTACTTCAGA 93622
QY 1571 AAGCATTTTTTTTCTTTTGAACCTTAAGCTTCTATATATAAGACAGACAGATTCCA 1630
      |||
Db 93621 AAGCATTTTTTTTCTTTTGAACCTTAAGCTTCTATATATAAGACAGACAGATTCCA 93562
QY 1631 CATTTTATACATGAGATCTTCTTGTGTGATACACAGATTGACTGATCCCTTAA 1690
      |||
Db 93561 CATTTTATACATGAGATCTTCTTGTGTGATACACAGATTGACTGATCCCTTAA 93503
QY 1691 AAGAGTTTATGTCCCTGACTGCTGCTAAATATATCTAATTTCCAGATGCTTTGTAGA 1750
      |||
Db 93502 AAGAGTTTATGTCCCTGACTGCTGCTAAATATCTAATTTCCAGATGCTTTGTAGA 93443
QY 1751 TGACGTGAGATTTTGAGGACACATATGGGATTTGAGTGTGAGTATGAGGAGAAA 1810
      |||
Db 93442 TGACGTGAGATTTTGAGGACACATATGGGATTTGAGTGTGAGTATGAGGAGAAA 93383
QY 1811 GGGCCATCTCCATGAGATGATTAAGTGAACCAACTAGTTCTCGAATTTCTACAGAAA 1870
      |||
Db 93382 GGGCCATCTCCATGAGATGATTAAGTGAACCAACTAGTTCTCGAATTTCTACAGAAA 93323
QY 1871 GGAGGGAATAGACTGAGAGAGCTGTGACATAGACTTGAAGACCAAGACTTTGAATTT 1930
      |||
Db 93322 GGAGGGAATAGACTGAGAGAGCTGTGACATAGACTTGAAGACCAAGACTTTGAATTT 93263
QY 1931 TGGGAGCTGCTCAGTGTGATGATTAATCACTGCTGCTTTTATTGAGTTACAAATCTA 1990
      |||
Db 93262 TGGGAGCTGCTCAGTGTGATGATTAATCACTGCTGCTTTTATTGAGTTACAAATCTA 93203
QY 1991 TATTTTATGAGTTTAATAAGAAAAAATTTCACAGA 2030
      |||
Db 93202 TATTTTATGAGTTTAATAAGAAAAAATTTCACAGA 93163

```

```

RESULT 11
AL354888 199863 bp DNA 1linear HTG 13-JUN-2001
LOCUS Homo sapiens chromosome 1 clone RP11-473p22, *** SEQUENCING IN
DEFINITION
ACCESSION AL354888 GI:9863692
VERSION AL354888.6 HTG: HTGS_PHASE1; HTGS_CANCELLED.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 199863)
AUTHORS Plumb,B.
TITLE Direct Submission

```

JOURNAL

Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 21, 2000 this sequence version replaced gi:9213379.

COMMENT

Center: Sanger Centre
Genome Center
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk

Center project name: BA473p22
Project Information
Center project name: BA473p22
Project Information

Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid, L08752; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 189476 bases at least Q40

Consensus quality: 193311 bases at least Q30
Consensus quality: 195510 bases at least Q20

Insert size: 201159; 6.2% error; sum-of-contigs
Quality coverage: 3.87x in Q20 bases; sum-of-contigs quality

coverage: 3.98x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently
consists of 16 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

```

1 2278: contig of 2278 bp in length
2279 2378: gap of 100 bp
2379 2278: gap of 100 bp
2379 2278: contig of 2035 bp in length
2274 22843: gap of 100 bp
22844 45183: contig of 22340 bp in length
45184 45283: gap of 100 bp
45284 49581: contig of 4298 bp in length
49582 49681: gap of 100 bp
49682 58258: contig of 8577 bp in length
58259 58358: gap of 100 bp
58359 64082: contig of 5724 bp in length
64083 64182: gap of 100 bp
64183 72764: contig of 8582 bp in length
72765 72864: gap of 100 bp
72865 86552: contig of 13688 bp in length
86553 86652: gap of 100 bp
86653 93481: contig of 6829 bp in length
93482 93581: gap of 100 bp
93582 101614: contig of 8033 bp in length
101615 101714: gap of 100 bp
101715 107836: contig of 6122 bp in length
107837 107936: gap of 100 bp
107937 144546: contig of 36610 bp in length
144547 144646: gap of 100 bp
144647 155892: contig of 11246 bp in length
155893 155992: gap of 100 bp
155993 167875: contig of 11883 bp in length
167876 167975: gap of 100 bp
167976 183411: contig of 15436 bp in length
183412 183511: gap of 100 bp
183512 199863: contig of 16352 bp in length.

```

FEATURES

source

1..199863
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-473p22"
/clone_1lb="RP11-11.2"

misc-feature

1..2278
/note="assembly_fragment:00765
fragment_chain:1
clone_end:T7

1149 2470: contig of 1322 bp in length
2471 2570: gap of 100 bp
2571 3643: contig of 1073 bp in length
3644 3743: gap of 100 bp
3744 4980: contig of 1237 bp in length
4981 5080: gap of 100 bp
5081 6352: contig of 1272 bp in length
6353 6452: gap of 100 bp
6453 7549: contig of 1097 bp in length
7550 8977: contig of 1328 bp in length
8978 9077: gap of 100 bp
9078 10103: contig of 1026 bp in length
10104 10203: gap of 100 bp
10204 11362: contig of 1159 bp in length
11363 11462: gap of 100 bp
11463 12821: contig of 1359 bp in length
12822 12921: gap of 100 bp
12922 14007: contig of 1086 bp in length
14008 14107: gap of 100 bp
14108 15283: contig of 1176 bp in length
15284 15383: gap of 100 bp
15384 16399: contig of 1016 bp in length
16400 16499: gap of 100 bp
16500 17804: contig of 1305 bp in length
17805 17904: gap of 100 bp
17905 19176: contig of 1272 bp in length
19177 19276: gap of 100 bp
19277 20384: contig of 1108 bp in length
20385 20484: gap of 100 bp
20485 21689: contig of 1205 bp in length
21690 21789: gap of 100 bp
21790 23191: contig of 1402 bp in length
23192 23291: gap of 100 bp
23292 24513: contig of 1222 bp in length
24514 24613: gap of 100 bp
24614 26135: contig of 1522 bp in length
26136 26235: gap of 100 bp
26236 27373: contig of 1138 bp in length
27374 27473: gap of 100 bp
27474 28723: contig of 1250 bp in length
28724 28823: gap of 100 bp
28824 30217: contig of 1394 bp in length
30218 30317: gap of 100 bp
30318 32056: contig of 1739 bp in length
32057 32156: gap of 100 bp
32157 33818: contig of 1662 bp in length
33819 33918: gap of 100 bp
33919 35156: contig of 1238 bp in length
35157 35256: gap of 100 bp
35257 36712: contig of 1456 bp in length
36713 36812: gap of 100 bp
36813 38379: contig of 1567 bp in length
38380 38479: gap of 100 bp
38480 40190: contig of 1711 bp in length
40191 40290: gap of 100 bp
40291 41780: contig of 1490 bp in length
41781 41880: gap of 100 bp
41881 44053: contig of 2173 bp in length
44054 44153: gap of 100 bp
44154 45474: contig of 1321 bp in length
45475 45574: gap of 100 bp
45575 46821: contig of 1247 bp in length
46822 46921: gap of 100 bp
46922 47302: contig of 381 bp in length
47303 47402: gap of 100 bp
47403 49323: contig of 1921 bp in length
49324 49423: gap of 100 bp
49424 51140: contig of 1717 bp in length
51141 51240: gap of 100 bp
51241 53099: contig of 1859 bp in length
53100 53199: gap of 100 bp
53200 54631: contig of 1432 bp in length

54632 54731: gap of 100 bp
54732 55751: contig of 1020 bp in length
55752 55851: gap of 100 bp
55852 57470: contig of 1619 bp in length
57471 57570: gap of 100 bp
57571 60182: contig of 2612 bp in length
60183 60282: gap of 100 bp
60283 63065: contig of 2783 bp in length
63066 63165: gap of 100 bp
63166 64443: contig of 1278 bp in length
64444 64543: gap of 100 bp
64544 66043: contig of 1500 bp in length
66044 66143: gap of 100 bp
66144 68623: contig of 2480 bp in length
68624 68723: gap of 100 bp
68724 70122: contig of 1399 bp in length
70123 70222: gap of 100 bp
70223 72407: contig of 2185 bp in length
72408 72507: gap of 100 bp
72508 74635: contig of 2128 bp in length
74636 74735: gap of 100 bp
74736 77554: contig of 2819 bp in length
77555 77654: gap of 100 bp
77655 80277: contig of 2623 bp in length
80278 80377: gap of 100 bp
80378 81798: contig of 1421 bp in length
81799 81898: gap of 100 bp
81899 84126: contig of 2228 bp in length
84127 84226: gap of 100 bp
84227 86794: contig of 2568 bp in length
86795 86894: gap of 100 bp
86895 90311: contig of 3417 bp in length
90312 90411: gap of 100 bp
90412 92255: contig of 1844 bp in length
92256 92355: gap of 100 bp
92356 94570: contig of 2215 bp in length
94571 94670: gap of 100 bp
94671 97163: contig of 2493 bp in length
97164 97263: gap of 100 bp
97264 99724: contig of 2461 bp in length
99725 99824: gap of 100 bp
99825 103151: contig of 3327 bp in length
103152 103251: gap of 100 bp
103252 106370: contig of 3119 bp in length
106371 106470: gap of 100 bp
106471 108886: contig of 2418 bp in length
108887 108988: gap of 100 bp
108989 111514: contig of 2526 bp in length
111515 111614: gap of 100 bp
111615 114360: contig of 2746 bp in length
114361 114460: gap of 100 bp
114461 118269: contig of 3809 bp in length
118270 118369: gap of 100 bp
118370 122359: contig of 3990 bp in length
122360 122459: gap of 100 bp
122460 125669: contig of 3210 bp in length
125670 125769: gap of 100 bp
125770 129306: contig of 3537 bp in length
129307 129406: gap of 100 bp
129407 132984: contig of 3578 bp in length
132985 133084: gap of 100 bp
133085 137360: contig of 4276 bp in length
137361 137460: gap of 100 bp
137461 140476: contig of 3016 bp in length
140477 140576: gap of 100 bp

Query Match 41.7% Score 850.4; DB 2; Length 200484;
Best Local Similarity 98.6% Pred. No. 6; E=211;
Matches 868; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
OY 1151 GATCGCATGGCGACACATTAATCCAAATCCACCTGGAACACGAGAGAAAGA 1210
DB 94041 GATTTCGTCTATACAGCGCATTAATCCAAATCCCACTTGAACCAACTGAGAAAGA 93982

Query Match	41.7%	Score 850.4;	DB 9;	Length 166496;	-
Best Local Similarity	98.6%	Pred. No. 6.6e-211;			
Matches 868; Conservative	0;	Mismatches 11;	Indels 1;	Gaps 1	

QY	1151	GAGCGGAATGGGCGGCACATAAAATCCAAATCCCACCTTGAACCACCTGAAGAAGA	1210
Db	82459	GATTTTGTCITTAATCAGGCCACATTAAATCCCAATCCCCACTTGAAGAAAAGA	824000
QY	1211	AGAAATTTGGAATCAGATGCTGTCCAACACCATAGAAAGTCAGAGTGTTCGCCAAGCTAT	1270
Db	82399	AGAAGATTGGAATCAGATGCTGTCCAACACCATAGAAAGTCAGAGTGTTCGCCAAGCAT	823400
QY	1271	AACAAGAACCTAAAGGAGGGATCCCCAGGCGAATGATCAAGAGCTGAATGCAAC	1330
Db	82339	AACAAGAACCTAAAGGAGGGATCCCCAGGCGAATGATCAAGAGCTGAATGCAAC	822800
QY	1331	GTTTAAGAGACATGTCACAGTGGCCTTTGGAAAAGTGTGTGGGATCCAGTTCAGAGGGAG	1390
Db	82279	GTTTAAGAGACATGTCACAGTGGCCTTTGGAAAAGTGTGTGGGATCCAGTTCAGAGGGAG	822200
QY	1391	GGGTATGTTTGTGCTCCCAAGTCAGGCGCAAGAGAGTCTATGCGGAATTCCTGCATAGCAG	1450
Db	82219	GGGTATGTTTGTGCTCCCAAGTCAGGCGCAAGAGAGTCTATGCGGAATTCCTGCATAGCAG	821600
QY	1451	AAAACTCCCAACCATTTTCTTTTGATGTGTTTTAAAGTCTCACGTTCTATATAGAA	1510
Db	82159	AAAACTCCCAACCATTTTCTTTTGATGTGTTTTAAAGTCTCACGTTCTATATAGAA	821000
QY	1511	ACAGAGGCTGTGTAGTCTGCTGTGTGGTGGTCAATGTGCTCGAAATGATGTAGTACAGA	1570
Db	82099	ACAGAGGCTGTGTAGTCTGCTGTGTGGTGGTCAATGTGCTCGAAATGATGTAGTACAGA	820400
QY	1571	AAGCATTTTTTTTTTCTTTGGAACCTTAAAGGCTCATATTATTTAAAGCAGACAGATTCCA	1630
Db	82039	AAGCATTTTTTTTTTCTTTGGAACCTTAAAGGCTCATATTATTTAAAGCAGACAGATTCCA	819800
QY	1631	CATTTTATACATGAGATCTTCTTTGTGTGAATACAGAGATTACGTGCATCCCTTAA	1690
Db	81979	CATTTTATACATGAGATCTTCTTTGTGTGAATACAGAGATTACGTGCATCCCTT-A	819210
QY	1691	AAGAAGTTTATGTCOCCTGAGATCTGTGGCTTAATAATATCTAATTTCCAGATGCTTTGTAGA	1750
Db	81920	AAGAAGTTTATGTCOCCTGAGATCTGTGGCTTAATAATATCTAATTTCCAGATGCTTTGTAGA	818610
QY	1751	TGACAGAGATATTTTGTGAGCACATATTGGGAGTCTCTGATTTGGATGTGATGGCAGGAAA	1810
Db	81860	TGACAGAGATATTTTGTGAGCACATATTGGGAGTCTCTGATTTGGATGTGATGGCAGGAAA	818010
QY	1811	GGGCGATCTCCATTTGAGATGATTAAAGTAGAACCAAACTAGTTCTCCGAATTTCTACAGAGAA	1870
Db	81800	GGGCGATCTCCATTTGAGATGATTAAAGTAGAACCAAACTAGTTCTCCGAATTTCTACAGAGAA	817410
QY	1871	GGAGGGATTCAGACTGAGAGAACTGTGACTATBGAACCTTGABAGACCAAAGACTTTGAATTT	1930
Db	81740	GGAGGGATTCAGACTGAGAGAACTGTGACTATBGAACCTTGABAGACCAAAGACTTTGAATTT	816810

RESULT 10	AC025002/c	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
	AC025002	200484 bp	DNA	Linear	HTG 14-JUN-2000			
		Homo sapiens chromosome 1 clone RP11-204L3 map 1,		*** SEQUENCING IN				
		PROGRESS *** 81 unordered pieces.						
	AC025002							
	AC025002.3	GI:8516086						
	HTG: HTGS_PPHASE1.							
	Homo sapiens.							
	Homo sapiens							
	Homo sapiens							

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE
1 (bases 1 to 200484)	Birren,B., Linton,L., Nusbaum,C. and Lander,E.	Homo sapiens chromosome 1, clone RP11-204L3	Unpublished	2 (bases 1 to 200484)

REFERENCE
AUTHORS

2 (bases 1 to 200484)

Björren, B., Linton, D., Nusbbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barua, N., Bastien, V., Beda, F., Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Campiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Deavelland, K., Dewar, K., Diaz, J. S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gaidyna, S., Ginde, S., Goyette, M., Graham, L., Howland, P., Iliev, I., Johnson, R., Jones, C., Kam, L., Karatas, A., Klein, J., Larocque, K., Lamaras, R., Landers, T., Lebochky, J., Levine, R., Lien, C., Liu, G., Locke, K., MacDonald, P., Marquis, N., McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McPheters, R., Melidim, J., Menues, L., Minova, C., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivari, T. M., Oliver, J., Peterson, K., Pierre, N., Pisanik, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Titrell, A., Travers, M., Trifillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE	Direct Submission
JOURNAL	Submitted (03-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Jun 14, 2000 this sequence version replaced gi:7770499. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information -----

Center project name: L7675
Center clone name: 204_L_3

* NOTE: This is a 'working draft' sequence. It currently

- * consists of 81 contigs. The true order of the pieces
- * is not known and their order in this sequence record is

- * arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.

- * This record will be updated with the finished sequence
- * as soon as it is available and the accession number will

* be preserved.
* 1 1048: contig of 1048 bp in length
* 1049 1148: gap of 100 bp

<http://bacpac.med.buffalo.edu/>
 VECTOR: pcYBAC2

This sequence is the entire insert of clone RPI-118J21. The true right end of clone RPI-11/L23 is at 38518 in this sequence. This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>.

FEATURES

Source
 1. 166496
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="p34.1-35.3"
 /clone="RPI-118J21"
 /clone_1lb="RPI-1"
 /complement(830..29381)
 /gene="BMP8"
 /complement(join(<830..1098,3622..3732,4242..4321,5153..5347,14838..14986,15379..15568,28682..29381))
 /gene="BMP8"
 /product="DJ118J21.1 (bone morphogenetic protein 8 (osteogenic protein 2))"
 /note="match: CDNAS: Em:S77477 Em:X51801 Em:M97017 Em:U39545 Em:U04566 Em:X80992 Em:M60316 Em:M60315 Em:M97016 Em:X56906 Em:AF100787 Em:U40034"
 /evidence="not_experimental"
 /complement(join(949..1098,3622..3732,4242..4321,5153..5347,14838..14986,15379..15568,28682..29015))
 /gene="BMP8"
 /note="match: proteins: Sw:P22004 Sw:P20722 Sw:P22003 Sw:P18075 Tr:O9YGH7 Sw:P23359 Sw:P55105 Sw:P49003 Sw:P34820 Sw:P34821"
 /codon_start=1
 /evidence="not_experimental"
 /product="DJ118J21.1 (bone morphogenetic protein 8 (osteogenic protein 2))"
 /protein_id="CAB75681.1"
 /db_xref="gi:7018292"
 /translation="MTALPGPLMLIGLALCALGGGPGLRPPGCPORRLCARRRDYORETLAVLGLGRPRAPAPASRLPASAPLEMLDLYAMAGDDDEDAPERRLGRA DLVASFVNMVERDALGHOEPMKEFRDLQIAPAGEVTAAEFRIVYPSIHILNRT LHVMEFQVVOBOSRESDFLDLOTLAGDEGMVLDVTAASDCMLKRRKDLGLRL YVEREDGSDVDPGLAGLIGRAPRSQGFVFPASPSPIRTPRAYPLRLRROPKKS NELPQANRIPGIPFDVYHSGRQVCRRLHLYVSFDCLMDLWVLAPOGTSYVCEGEC SFLDSCNATNHA1LOSVLVHMPDAVPAKACAPKTSASVLYDSSNNVTLKRR NMVVKACGCH"
 1655..1748
 /note="L2 repeat: matches 1691..1787 of consensus"
 1849..2281
 /note="L2 repeat: matches 2002..2438 of consensus"
 2764..2868
 /note="FLAM_C repeat: matches 1..133 of consensus"
 /complement(4214..4420)
 /gene="BMP8"
 /note="match: STS: Em:T63436"
 /complement(4252..4438)
 /gene="BMP8"
 /note="match: STS: Em:AA634039"
 5972..6029
 /note="MIR repeat: matches 19..75 of consensus"
 6293..6588
 /note="Aluub repeat: matches 1..307 of consensus"
 7787..7873
 /note="MIR repeat: matches 57..146 of consensus"
 /complement(8369..8876)
 /gene="BMP8"
 /note="match: GSS: Em:B51067"
 9109..9404
 /note="Aluub repeat: matches 1..296 of consensus"
 9427..9746
 /note="Aluub repeat: matches 1..312 of consensus"
 10075..10242

misc_feature
 10626..12048
 /note="Cpg Island"
 /evidence="not_experimental"
 12351..13402
 /note="L2 repeat: matches 957..2141 of consensus"
 13474..13591
 /note="L2 repeat: matches 178..295 of consensus"
 13596..14185
 /note="Aluuo/FRAM repeat: matches 178..295 of consensus"
 14186..14425
 /note="L1MC3 repeat: matches 7071..7691 of consensus"
 14457..14572
 /note="L2 repeat: matches 2129..2417 of consensus"
 15137..15215
 /note="MIR repeat: matches 27..147 of consensus"
 16285..16510
 /note="MIR repeat: matches 47..137 of consensus"
 17313..17597
 /note="L2 repeat: matches 2361..2586 of consensus"
 20320..20782
 /note="L2 repeat: matches 2445..2745 of consensus"
 20785..20932
 /note="MTRIF repeat: matches 3..541 of consensus"
 21140..21425
 /note="MIR repeat: matches 22..175 of consensus"
 21927..22073
 /note="Alusx repeat: matches 1..282 of consensus"
 22185..22408
 /note="MIR repeat: matches 97..262 of consensus"
 22409..22704
 /note="LTR16A repeat: matches 116..325 of consensus"
 22705..22759
 /note="Alusg repeat: matches 5..306 of consensus"
 22829..22898
 /note="LTR16A repeat: matches 60..116 of consensus"
 22952..22989
 /note="MIR repeat: matches 5..71 of consensus"
 22990..23298
 /note="MIR repeat: matches 227..261 of consensus"
 23299..23381
 /note="Alusq repeat: matches 1..308 of consensus"
 23401..23576
 /note="MIR repeat: matches 132..227 of consensus"
 23600..23742
 /note="Alusx repeat: matches 51..223 of consensus"
 24157..24451
 /note="MIR repeat: matches 8..153 of consensus"
 24618..24684
 /note="Aluio repeat: matches 2..294 of consensus"
 24753..24891
 /note="MIR repeat: matches 80..147 of consensus"
 25018..25159
 /note="MIR repeat: matches 50..199 of consensus"
 26215..26275
 /note="L2 repeat: matches 2606..2750 of consensus"
 26604..26643
 /note="Alusp repeat: matches 143..195 of consensus"
 26930..27021
 /note="MIR repeat: matches 80..119 of consensus"
 27580..27733
 /note="MIR repeat: matches 81..179 of consensus"
 28059..28151
 /note="MIR repeat: matches 37..197 of consensus"
 28542..30030
 /note="L2 repeat: matches 2598..2703 of consensus"
 /note="Cpg Island"
 29451..29584
 /evidence="not_experimental"
 30126..30262
 /note="67 copies 2 mer gg 598 conserved"
 30283..30460
 /note="L2 repeat: matches 2350..2475 of consensus"
 /note="MIR repeat: matches 60..241 of consensus"

Web site: <http://www.bpsc.bcm.tmc.edu/cdna/>
 Contact: amgedbcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/BLN! at: <http://Image.lnl.gov>
 Series: IRAP Plate: 41 Row: 1 Column: 20.
 Location/Qualifiers
 1. 2045
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:30541 IMAGE:5042856"
 /tissue.type="Kidney, normal, 5 month old male mouse."
 /clone.lib="NCI GAP_Kid14"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 406. 1386
 /codon_start=1
 /product="Riken cDNA 2310075G14 gene"
 /protein_id="AAH19812.1"
 /db_xref="GI:18044186"
 /db_xref="LocusID:66966"
 /translation="MGTGKVVYDRKVELEKEDGHEIARRLSQVDEPMAKLHPDRKRY
 ARSLQFEETGISHSEFIRHROAEEGGSLGPRPAPCILMLHADOAVLDERLDKPE
 VDMLAAGLLEELRGPHRYNLKINSISNOYHGISTGKFEHFLTEKCTPE
 TSNLDLKGKIFALKQVTRKRYARKNRKRYKRFESRPQSPVYGLVSDYSKDEEST
 LEPNLNVOSTFGHKRTPAMPVKRAVSEKRSYHMDLDRITIGREAAHLKSS
 SHLQLRKRRLDIDAVSATOSNSPSCPDERTEGSSGQNOELKASV"

BASE COUNT 570 a 453 c 541 g 481 t
 ORIGIN

Query Match	54.8%	Score 1119.4	DB 10	Length 2045
Best Local Similarity	75.8%	Prod. No. 1.le-281		
Matches 1584	Conservative	0	Mismatches 381	Indels 126; Gaps 11
Oy	29	GCACGAGCAGTCCCTGCTGGGCGAGTGGGCTGCAGAGGCGCTGCACGAGACCTTACCCTTGTGA	88	
Dd	1	GCACGAGCAGTCCCTGCTGAGCTCTGGATTCCGAGGCTCGCGCAACACTGCCACTTGTGA	60	
Oy	89	GTGATTCTCGGGGCGACGGGCGACGGCAAAATCCACGCTGGCGCTTGCAGCTAGGCGACGGG	148	
Dd	61	GTGATTCTCGGGGCTACTGGGACTGGCAAGTCCACCTCGGCTGTGCAAGTTAGCCAGCGG	120	
Oy	149	CTCGGCGGTGAGATGCTCAGCGCTTACATTCATGCAAGGTCTATGAAGGCGCTACACATCATC	208	
Dd	121	CTCGGCGGCGAATGTGTAGCGGCCACATCCATGCAAGGCTTATGAAGGCGTACATCATC	180	
Oy	209	ACCACACAGGTTTCCGCCAAGAGAGAGAAATCTCGCGGCGACACATGATCAGCTTGTG	268	
Dd	181	ACCATTAAAGGTTCTGCCCCAAGAGAGAGATGTCCACACCATGATCAGCTTGTG	240	
Oy	269	GATCCTCTTGACCAATTTACACAGTGTGTGACTTCAGAAATATGAGCAACTGCTGTGAT	328	
Dd	241	GATCCTCTTGAGACAGGTACACGCTGGTGGACTTTAGGAACAAGCAACGCGCTGAT	300	
Oy	329	GAAATATATTTTGGCCGAGACAAATTTCTTATTTGTTGGGAGGAACCAATTTATTAAT	388	
Dd	301	GAAATATATTTTGTCTGAGACAAATTTCCATTTGTGTTGGGAGGAACCAATTTATTAAT	360	
Oy	389	GAAATCTGCTGTGAAAGTTCTTGTCAATATACAGCCCGACAGATATGGCGCTAGAAA	448	
Dd	361	GAAATCTTCTGTGAAAGTTCTTGTCAATATACAGCCCGACAGATATGGCGCTAGAAA	420	
Oy	449	GTGATTGACCGAAAAGTGTGAAAAAGAGAGATGTTCTTGTACTTTCACAAACGCTGA	508	
Dd	421	GTGATTGATCGAAAGTGTGAAAAAGAGATGTTCTTGTACTTTCACAAACGCTGA	480	
Oy	509	AGCCAGGTGAGACCGAAGATGCTGCCAAGCTGCATCCACATGACAAACGCAAAATGGCC	568	
Dd	481	AGCCAGGTGAGACCGAAGATGCTGCCAAGCTGCATCCACATGACAAACGCAAAATGGCC	540	

KEYWORDS oligo capling; fis (full insert sequence).

SOURCE Homo sapiens colon cDNA to mRNA, clone lib:COL clone:COL01371.

ORGANISM Homo sapiens

REFERENCE 1
Kawabata,A., Hiki,J.T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shihabara,T., Tanaka,Y., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1749)

TITLE
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shihabara,T., Tanaka,Y. and Nakamura,Y.
Direct Submission
Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan: cDNA full insert sequencing; Research Association for Biotechnology: cDNA library construction; 5'-3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
Location/Qualifiers
1. 1749

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL01371"
/tissue_type="colon"
/clone_lib="COL"
/note="cloning vector: pME18SFL3"

BASE COUNT 541 a 337 c 408 g 463 t

ORIGIN

Query Match 76.3%; Score 1556.8; DB 9; Length 1749;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 7; Indels 79; Gaps 2;

OY 374 ACCAATTATTAACCTGTAACCTGCTGGAAGTCTTGTCAATCCAGCCCGCAGAG 433
DB 1 ACCAATTATTAACCTGTAACCTGCTGGAAGTCTTGTCAATCCAGCCCGCAGAG 60

OY 434 ATGGGCACTGAGAAAGTGAATGACCGAAAGTGGAGCTTGAAGAGAGATGCTTGTGA 493
DB 61 ATGGGCACTGAGAAAGTGAATGACCGAAAGTGGAGCTTGAAGAGAGATGCTTGTGA 120

OY 494 CTTCACAAAGCCTTAAGCCAGGTGACCCAGAAATGGCTGCCAAGCTGCATCCATGAC 553
DB 121 CTTCACAAAGCCTTAAGCCAGGTGACCCAGAAATGGCTGCCAAGCTGCATCCATGAC 180

OY 554 AAAGCGAAGTGGCCAGAGCTGCAAGTTTGAAGAAACAGAGATCTCTCATGTGAA 613
DB 181 AAAGCGAAGTGGCCAGAGCTGCAAGTTTGAAGAAACAGAGATCTCTCATGTGAA 240

OY 614 TTCTCCATGTCACATACGGAAGAGTGTGTCCTCCCTTGAGAGTCTCTGGAATTG 673
DB 241 TTCTCCATGTCACATACGGAAGAGTGTGTCCTCCCTTGAGAGTCTCTGGAATTG 300

OY 674 TCTAACCTTGACCTCTTGCTTCATGCTGACAGCAGCTGTAGATGAGCGCTTGAT 733
DB 301 TCTAACCTTGACCTCTTGCTTCATGCTGACAGCAGCTGTAGATGAGCGCTTGAT 360

OY 734 AAGAGGTGATGACATGCTGCTGGGCTTGGAGGAAGTGAAGATTTCACAGA 793
DB 361 AAGAGGTGATGACATGCTGCTGGGCTTGGAGGAAGTGAAGATTTCACAGA 420

OY 794 CGGTATATCAGAGAATGTTTGGAAAATAGCCAGGACTATCAATGATATCTTCCAA 853
DB 421 CGGTATATCAGAGAATGTTTGGAAAATAGCCAGGACTATCAATGATATCTTCCAA 480

OY 854 TCAATTGGCTTCAAGAAATTCACGAGTACCTATCATGAGGAAATGCAACTGAG 913
DB 481 TCAATTGGCTTCAAGAAATTCACGAGTACCTATCATGAGGAAATGCAACTGAG 540

OY 914 ACTAGTACCAAGCTTTAAAGAAAG----- 939
DB 541 ACTAGTACCAAGCTTTAAAGAAAGTATTGAGGCTCTGAACAACTAATAAGACATAT 600

OY 940 -----ACTGGTCCCATGTC 955
DB 601 GCCCGAACAACCGATGGTTAAACCCGTTTTTGGACAGACCTGTCATTTGTC 660

OY 956 CCCCCTGTCTAGGCTTGAAGATATGATGTCGAAAGTGGAGAGATGCTTTTGA 1015
DB 661 CCCCCTGTCTAGGCTTGAAGATATGATGTCGAAAGTGGAGAGATGCTTTTGA 720

OY 1016 CCGGCTTGAATCGTGCAGAACTTCATCCAGGCGCAAGAGCTCAGCCATCCAAAT 1075
DB 721 CCGGCTTGAATCGTGCAGAACTTCATCCAGGCGCAAGAGCTCAGCCATCCAAAT 780

OY 1076 AAGATGCCATATCAATGAGCTGAGAACAGAGAAGTTATCCCTGTGACTGTGAT 1135
DB 781 AAGATGCCATATCAATGAGCTGAGAACAGAGAAGTTATCCCTGTGACTGTGAT 840

OY 1136 CGAATCATATGGGATCGGATGCGAATGGCGACGACATTAATCCAAATCCACTGAA 1195
DB 841 CGAATCATATGGGATCGGATGCGAATGGCGACGACATTAATCCAAATCCACTGAA 900

OY 1196 CAATGAGAAAGAAAGATGAGATGAGATCGATGCTCAACCAATAGAAATCGAGT 1255
DB 901 CAATGAGAAAGAAAGATGAGATGAGATCGATGCTCAACCAATAGAAATCGAGT 960

OY 1256 GTTTCGCCAGACTATTAACAAGAACTTAAGGAGAGGATCCCGAGGCGAATGATCAA 1315
DB 961 GTTTCGCCAGACTATTAACAAGAACTTAAGGAGAGGATCCCGAGGCGAATGATCAA 1020

OY 1316 GACCTGAATGACGCGTTTAAAGACATGTCAGTGGCTTTGGAAGTGTGGGATC 1375
DB 1021 GACCTGAATGACGCGTTTAAAGACATGTCAGTGGCTTTGGAAGTGTGGGATC 1080

OY 1376 CAGTTGAGGAGGAGGGGTGTTTGTCTCCAGTCCGCAAGAGAGTCTATGCGGAA 1435
DB 1081 CAGTTGAGGAGGAGGGGTGTTTGTCTCCAGTCCGCAAGAGAGTCTATGCGGAA 1140

OY 1436 TTCTCTGATAGCAGAAAGCTCCACCATTTCTTTGATGTGTTTAAAGTCTCAG 1495
DB 1141 TTCTCTGATAGCAGAAAGCTCCACCATTTCTTTGATGTGTTTAAAGTCTCAG 1200

OY 1496 TTCTCTATATTAACAAACAGCAGCTCTTGTGAGCTCTGTGTGATGTCTGGAA 1555
DB 1201 TTCTCTATATTAACAAACAGCAGCTCTTGTGAGCTCTGTGTGATGTCTGGAA 1260

OY 1556 TGAATGATGACAGAAAGCAATTTTCTTGAACCTTAAGGTTCTATTATTTAAA 1615
DB 1261 TGAATGATGACAGAAAGCAATTTTCTTGAACCTTAAGGTTCTATTATTTAAA 1320

OY 1616 GCAGCAGAGATTCACATTTTATACATGAGATCTTCTTGTGTGTAATACAGAGAT 1675
DB 1321 GCAGCAGAGATTCACATTTTATACATGAGATCTTCTTGTGTGTAATACAGAGAT 1380

OY 1676 ACTGCATCCCTTTAAAGAAAGTTTATGTCCGACCTGTGGCTTAAATATCTAATTTCC 1735
DB 1381 ACTGCATCCCTTTAAAGAAAGTTTATGTCCGACCTGTGGCTTAAATATCTAATTTCC 1439

OY 1736 AGATGCTTTGTATGATGATGAGTATTTTGAAGCCATATTTGGAGTCTGATTTGA 1795
DB 1440 AGATGCTTTGTATGATGATGAGTATTTTGAAGCCATATTTGGAGTCTGATTTGA 1499

OY 1796 GTAATGCGAGAGAGGCGCATCTCCATTGATGATTAAGTAAACCAACTAGTTCTCG 1855
DB 1500 GTAATGCGAGAGAGGCGCATCTCCATTGATGATTAAGTAAACCAACTAGTTCTCA 1559

OY 1856 GAATTTACAGAGAGAGGGAATCAGACTGAGAGAGCTGTGACATTAAGACTTGAAGACC 1915

```

"note="unnamed protein product"
/codon_start=1
/protein_id="BAA90923.1"
/db_xref="GI:7019915"
/translaton="MGTETKVIDRKEVLEKEDGLVLHKRLSQVDDEMAKALPHDHRKRY
ARSLQVPEERGISSEFLRQHTKEEGSGAGGGLFSPNCILMTHADQVAVDELDRK
VDMLAGLGLLEELRDLDFHRRYRNOKWNSNSDQYHGISTQIGKEFHYLITEKCTLE
TSNQLKLGKITALQVTKRYARKQNRNRYKKNFSLRPGPIYVPYVGLKSYDSVKSEESY
LEPDLVQSFTQSGKRPATPIFKMRYVNAEKKRSYTHLCDLDRILLIDGREAAHIKSK
SHLNLQKRRRLSDAVNTIESQSVSDSPNSDKEPKESPGQNDQELCSV"
541 a 337 c 408 g 463 t

```

Query Match	76.3%	Score 1556.8	DB 9	Length 1749
Best Local Similarity	95.1%	Pred. No. 0		
Matches 1600	Conservative	0	Mismatches 7	Indels 79
				Gaps 2

OY	374	ACAAATTTATCATGTGAATCTCTGCTCGAAGAGTTCTTGCAATACCAAGCCCAAGAG	433
Db	1	ACCAATTTATCATGTGAATCTCTGCTCGAAGAGTTCTTGCAATACCAAGCCCAAGAG	60
OY	434	ATGGGCACTGAGAAAGTATGACCGAAAGTGAAGTGAAGTGAAGAGAGATGCTTTGA	493
Db	61	ATGGGCACTGAGAAAGTATGACCGAAAGTGAAGTGAAGTGAAGAGATGCTTTGA	120
OY	494	CTTCACAAACGCTTAAGCAGGTGGACCCAGAAATGCGTCCCAAGCTGATCACAATGAC	553
Db	121	CTTCACAAACGCTTAAGCAGGTGGACCCAGAAATGCGTCCCAAGCTGATCACAATGAC	180
OY	554	AAACGCAAGTGGCCAGGAGGCTGCAAGCTTTTGAAGAAACGAAATCCTCATGTGTA	613
Db	181	AAACGCAAGTGGCCAGGAGGCTGCAAGCTTTTGAAGAAACGAAATCCTCATGTGTA	240
OY	614	TTTCTCCATCGTCAACATACGGAAGAAGTGGTGGTCCCTTGGAAGTCCCTGGAAGTTC	673
Db	241	TTTCTCCATCGTCAACATACGGAAGAAGTGGTGGTCCCTTGGAAGTCCCTGGAAGTTC	300
OY	674	TCTAACCTTGATCCTTTGGCTTCATGCTGACCGAGCAAGTTCTAGATGAGGCGTTGGAT	733
Db	301	TCTAACCTTGATCCTTTGGCTTCATGCTGACCGAGCAAGTTCTAGATGAGGCGTTGGAT	360
OY	734	AAGAGGCGTGAATGACATGCTTGCTGCTGGGCTCTTGAGAGAACTAAGAGATTTTCACAG	793
Db	361	AAGAGGCGTGAATGACATGCTTGCTGCTGGGCTCTTGAGAGAACTAAGAGATTTTCACAG	420
OY	794	CGCTATATTCAGAGAAATGTTTCGAAAAATACCGACGACTATCAACATGATGTTCTCCAA	853
Db	421	CGCTATATTCAGAGAAATGTTTCGAAAAATACCGACGACTATCAACATGATGTTCTCCAA	480
OY	854	TCAATTGCTTCAAGGAATTCACAGATACCGATACATGAGGGAAATGCAACTGTGGAG	913
Db	481	TCAATTGCTTCAAGGAATTCACAGATACCGATACATGAGGGAAATGCAACTGTGGAG	540
OY	914	ACTAGTAAACGAGCTTCTTAAGAAAAAG-----	939
Db	541	ACTAGTAAACGAGCTTCTTAAGAAAAAG-----	600
OY	940	-----ACCTGGTCCATGTGC	955
Db	601	-----ACCTGGTCCATGTGC	660
OY	956	CCCCCTGTCTAATGGCTTAGAGGATACGATGCTCGAAGTGGGAGAGATCTGCTTTGAA	1015
Db	661	CCCCCTGTCTAATGGCTTAGAGGATACGATGCTCGAAGTGGGAGAGATCTGCTTTGAA	720
OY	1016	CTGTGCTTGAATTCGTCAAAAGTTTCATCCAGGGCCCAAGCCCTACACCACTCCAAATA	1075
Db	721	CTGTGCTTGAATTCGTCAAAAGTTTCATCCAGGGCCCAAGCCCTACACCACTCCAAATA	780
OY	1076	AAGATGCCATACATGAAGCTGAGAACAGAGAGACTATACCTGTGTGACCTCTGTGAT	1135
Db	781	AAGATGCCATACATGAAGCTGAGAACAGAGAGACTATACCTGTGTGACCTCTGTGAT	840

QY	1136	CGAATCATCATTTGGGGATGCGCAATGGGCGACGCGCACATATAATCCAAATCCCACTTGAAAC	1195
Db	841	CGAAATCATCATTTGGGGATGCGCAATGGGCGACGCGCACATATAATCCCAATCCCACTTGAAAC	900
QY	1196	CAACGTAAGAAAGAAGAATTTGACATCTAGATGCTGCTCAACACATATGAAGTCAGAGT	1255
Db	901	CAACGTAAGAAAGAAGAATTTGACATCTAGATGCTGCTCAACACATATGAAGTCAGAGT	960
QY	1256	GTTCCTCCAGACATATAACAAAGAACCTAAAGGGAAGGATCCCAAGGGCGAGTAATGACAA	1315
Db	961	GTTCCTCCAGACATATAACAAAGAACCTAAAGGGAAGGATCCCAAGGGCGAGTAATGACAA	1020
QY	1316	GAGCTGAATATGACAGCGCTTTAAGACACATGTCCAGTGGCGCTTTGGAAGGTGTGGGGATC	1375
Db	1021	GAGCTGAATATGACAGCGCTTTAAGACACATGTCCAGTGGCGCTTTGGAAGGTGTGGGGATC	1080
QY	1376	CAGTTCAGAGGAGGAGGGATGTGTTGTCTCCAGTCCGAGTGGGCAAGAGAGTATCGGAA	1435
Db	1081	CAGTTCAGAGGAGGAGGGATGTGTTGTCTCCAGTCCGAGTGGGCAAGAGAGTATCGGAA	1140
QY	1436	TTCTCTGCATAGCAGAAAAAGCTCCACCACTTTCTTTTGATGTGTTTAAAGTCTCAGC	1495
Db	1141	TTCTCTGCATAGCAGAAAAAGCTCCACCACTTTCTTTTGATGTGTTTAAAGTCTCAGC	1200
QY	1496	TTCTCTATTAATATGAAGACAGCAGTCTGTGTCAGCTCCTGTGTGCTGTATGTCTCGGAAA	1555
Db	1201	TTCTCTATTAATATGAAGACAGCAGTCTGTGTCAGCTCCTGTGTGCTGTATGTCTCGGAAA	1260
QY	1556	TGATGTAGTTCAGGAAGACATTTTCTTTTCTTTGAACCTTAAAGCTTATATTATTAATA	1615
Db	1261	TGATGTAGTTCAGGAAGACATTTTCTTTTCTTTGAACCTTAAAGCTTATATTATTAATA	1320
QY	1616	GCAGCAGACAGATTCACATTTTATATACATGAGAGATCTCTTTGTGTGATACACAGATATG	1675
Db	1321	GCAGCAGACAGATTCACATTTTATACATGAGAGATCTCTTTGTGTGATACACAGATATG	1380
QY	1676	ACTGCATCCCTTTAAAAAGAGTTTATATGTCCTGACTCTGGCTAAATAATATCATATTCC	1735
Db	1381	ACTGCATCCCTTT - AAAAGAGTTTATATGTCCTGACTCTGGCTAAATAATATCATATTCC	1439
QY	1736	AGATCCTTTTGTAGATGACTGAAGTATTTGTGAGCCACATATTTGGGAGTTCAGATTGGA	1795
Db	1440	AGATCCTTTTGTAGATGACTGAAGTATTTGTGAGCCACATATTTGGGAGTTCAGATTGGA	1499
QY	1796	GTGAATGGCAGAAAGGCCATCTCCATTTGAGATGATTAAGTGAACCAAACTAGTTCTCG	1855
Db	1500	GTGAATGGCAGAAAGGCCATCTCCATTTGAGATGATTAAGTGAACCAAACTAGTTCTCA	1559
QY	1856	GAAATCTACAGAGAGAGAGGAAATACAGACTGAGGAAGCTGTGACATATAGGACTTGAAAGCC	1915
Db	1560	GAAATCTACAGAGAGAGGAGGAAATACAGACTGAGGAAGCTGTGACATATAGGACTTGAAAGCC	1619
QY	1916	AAAGACTTTGAATTTGGAGCGTCTCATGTGTGAGTTATTTACATCGTCTGCTTTCTAT	1975
Db	1620	AAAGACTTTGAATTTGGAGCGTCTCATGTGTGAGTTATTTACATCGTCTGCTTTCTAT	1679
QY	1976	TGAGTTACAAATCTATATTTTTATTTGAAGTTTAAATAAGAAAAAATTTACACAAGAAAAA	2035
Db	1680	TGAGTTACAAATCTATATTTTTATTTGAAGTTTAAATAAGAAAAAATTTAAAAAATAAAAAA	1739
QY	2036	AAAAAA 2041	
Db	1740	AAAAAA 1745	

LOCUS	AK074222	1749 bp	mRNA	linear	PRI 15-FEB-2002
DEFINITION	Homo sapiens CDNA FLJ23642.f1s, clone COL01371, highly similar to Homo sapiens tRNA isopentenylpyrophosphate transferase precursor tRNA.				
ACCESSION	AK074222				
VERSION	AK074222.1	GI:18676765			

```

QY 661 TCCTGAGAGTCTCTAACCCCTGCATCCCTGGCTTCAGCTGACAGGACAGTCTAGA 720
    |||
Db 665 TCCTGAGAGTCTCTAACCCCTGCATCCCTGGCTTCAGCTGACAGGACAGTCTAGA 724
QY 721 TGAGCCCTGGATAGAGGGTGGATGACATGCTGCTGGGCTTTGGAGAGTAAG 780
    |||
Db 725 TGAGCCCTGGATAGAGGGTGGATGACATGCTGCTGGGCTTTGGAGAGTAAG 784
QY 781 AGATTTTTCACAGACGCTATATATCAGAAAGATGTTTCGGAAAAAT- 823
    |||
Db 785 AGATTTTTCACAGACGCTATATATCAGAAAGATGTTTCGGAAAAATGAAATGGCTCAAC 844
QY 824 - 823
Db 845 TAGGTCATCTCCAAAGCCCTGCTAGCTTTAGAGCAATGGCTTATCTGTGATACC 904
QY 824 - AGCCAGAGACTATACATGATGATCTTCACATCAATGGCTTCAAGAAAT 873
    |||
Db 905 TCCTGTTCCAGCCAGGACATATACATGATGATCTTCACATCAATGGCTTCAAGAAAT 964
QY 874 TCACGAGTACCTGATCAGTGAAGGAAATGACACTGGAGACTATGAACAGCTTCTAAA 933
    |||
Db 965 TCACGAGTACCTGATCAGTGAAGGAAATGACACTGGAGACTATGAACAGCTTCTAAA 1024
QY 934 GAAAG- 938
    |||
Db 1025 GAAAGGATTTGAGGCTCTGAACAAGTAACTAAGAGATATGCCGGAACAAACCGATG 1084
QY 939 - GACCTGGTCCCATTTGCTCCCGCTGCTATGAGTCTAGA 975
    |||
Db 1085 GATTAAAAACCGTTTGTGAGCAGACCTGGTCCCATTTGCTCCCGCTGCTATGAGTCTAGA 1144
QY 976 GGTATCTGATGCTCTGAGAGTGGAGAGAGTGTCTTGAACCTGCTCTTGAATCTGCA 1035
    |||
Db 1145 GGTATCTGATGCTCTGAGAGTGGAGAGAGTGTCTTGAACCTGCTCTTGAATCTGCA 1204
QY 1036 AAGTTTCATCCAGGCGCACAGGCTTACAGGCTCCATTAAGATGCCATATGAGC 1095
    |||
Db 1205 AAGTTTCATCCAGGCGCACAGGCTTACAGGCTCCATTAAGATGCCATATGAGC 1264
QY 1096 TGAACAACAAGAAATATATCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1155
    |||
Db 1265 TGAACAACAAGAAATATATCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1324
QY 1156 CGATGGGCGACGACATATAATCCAAATCCACTTGAACCACTGAAGAAAGAGAAAG 1215
    |||
Db 1325 CGATGGGCGACGACATATAATCCAAATCCACTTGAACCACTGAAGAAAGAGAAAG 1384
QY 1216 ATTGAGACTCAGATGCTGTCAACACATAGAAAGTGTGTTCCCGACATATAAGAA 1275
    |||
Db 1385 ATTGAGACTCAGATGCTGTCAACACATAGAAAGTGTGTTCCCGACATATAAGAA 1444
QY 1276 AGAAGCTAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1335
    |||
Db 1445 AGAAGCTAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1504
QY 1336 AGAGACATGTCCAGTGGCTTTGAGAAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGG 1395
    |||
Db 1505 AGAGACATGTCCAGTGGCTTTGAGAAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGG 1564
QY 1396 TGTGTTGCTCCAGTCTGGGCAAAAGAGTGTCTATGCGGAATTTCTCTCATAGCAGAAAG 1455
    |||
Db 1565 TGTGTTGCTCCAGTCTGGGCAAAAGAGTGTCTATGCGGAATTTCTCTCATAGCAGAAAG 1624
QY 1456 CTCACACATTTTCTTTGATGTTGTTTAAAGTCTCAGCTTCTCTATATAGAAACAGC 1515
    |||
Db 1625 CTCACACATTTTCTTTGATGTTGTTTAAAGTCTCAGCTTCTCTATATAGAAACAGC 1684
QY 1516 AGGTCTTGTGAGCTCTTGTGTGCTGATGTCTCTGAGAAATGATGTAGTTCAGAGAAAGCA 1575
    |||
Db 1685 AGGTCTTGTGAGCTCTTGTGTGCTGATGTCTCTGAGAAATGATGTAGTTCAGAGAAAGCA 1744
QY 1576 TTTTCTTTTCTTGAACCTTAAAGGTTCTATATATTAAGAGCAGACAGATTCACACATTT 1635

```

```

Db 1745 TTTTCTTTTCTTTGTAACCTTAAAGTCTATATTTAAAGCAGCAGATTCACATTT 1804
    |||
QY 1636 TTATACATGAGATCTCTTGTGTGTAATACAGAGATGACTGCATCCCTTTAAAGAA 1695
    |||
Db 1805 TTATACATGAGATCTCTTGTGTGTAATACAGAGATGACTGCATCCCTTT-AAAGAA 1863
QY 1696 GTTTTATGTCCTGACTCTGGCTAAATATCTAATTTTCCAGATGCTTTTGTGATGACT 1755
    |||
Db 1864 GTTTTATGTCCTGACTCTGGCTAAATATCTAATTTTCCAGATGCTTTTGTGATGACT 1923
QY 1756 GAAGTATTTGAGCCACATATTGGGAGTCTGTGATTTGAGTAAAGCAGAGAGGCC 1815
    |||
Db 1924 GAAGTATTTGAGCCACATATTGGGAGTCTGTGATTTGAGTAAAGCAGAGAGGCC 1983
QY 1816 ATCTCCATGAGATGATTAAGTAAAGCAAACTAGTTCTCGAATTTACAGAGAGAGG 1875
    |||
Db 1984 ATCTCCATGAGATGATTAAGTAAAGCAAACTAGTTCTCGAATTTACAGAGAGAGG 2043
QY 1876 GAATCAGACTGAGAGAGCTGTGACATGACACTTGAAGACCAGAACTTTGAATTTGCCA 1935
    |||
Db 2044 GAATCAGACTGAGAGAGCTGTGACATGACACTTGAAGACCAGAACTTTGAATTTGCCA 2103
QY 1936 GCTGCTCATGCTGAGTATATACAGTCTGCTCTTCTATGAGTAAATGATATTT 1995
    |||
Db 2104 GCTGCTCATGCTGAGTATATACAGTCTGCTCTTCTATGAGTAAATGATATTT 2163
QY 1996 TTATTAAGTTTAAATTAAGAAAAATTTACAGAAAAAATAAAAAA 2041
    |||
Db 2164 TTATTAAGTTTAAATTAAGAAAAATTTTAAAAAATAAAAAA 2209
    |||

RESULT 4
AK000068 1749 bp mRNA linear PRI 22-FEB-2000
LOCUS AK000068
DEFINITION Homo sapiens cDNA FLJ20061 fis, clone COL01383.
ACCESSION AK000068
VERSION AK000068.1 GI:7019914
KEYWORDS oligo capping; fis (full insert sequence);
SOURCE Homo sapiens colon cDNA clone, clone_11b:COL clone:COL01383.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites)
Kawabata,A., Hiki,I., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1749)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp,
Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing: Research Association for Biotechnology: cDNA library
construction, 5'-63'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
1. 1749
Location/Qualifiers
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL01383"
/tissue_type="colon"
/clone_lib="COL"
/note="cloning vector pME18SFL3"
CDS
61..1041

```

	db_xref="taxon:9606"	/note="extension of sequence by 5' RACE"
precursor_RNA	1. .2216	/note="Incompletely processed mRNA"
CDS	join(15. .829,917. .1505)	/note="lprase"
	/codon_start=1	/product="tRNA Isopentenyl pyrophosphate transferase"
	/protein_id="AA031324.1"	
	/db_xref="GI:1137965"	
	/translation="MASVAAARAAPVPGSGRLGRLTLPYLVIIGATGTGKTALALG ORGLGEIVSADSMQVGEGLDIFKFKVASQORLIRHMI SFVDPLVINYVDDRBRRA TALIEDIFARDIPIVYGNTNYIESLILMYLNTKROEGTEVIDRKELEDGI VLKRIISOVDEMAAKLHPHDKRKVASLQVFEPTGI SISEPFLHROHREGGGLGRR LKTSNPCIWLHLDQAVLEDERLDRVDMILAAGLELRDPHRYRNRKNVSENSDYQ HGIFQSIGKEFEHEYLITTEKCTLETSNOLKIGLIALKQVTKYARKONVNRRL SRGRPIVPYVGLSEVSDSVKWEESVLEPALEIYVSFIOGNRPATPKMPENENR SYHLCDLCRIIIGDREMAHIKSKSHLNOLKRRRLSDAVNITIEOSVSPDNKEP KENGSPONDQELKCSV"	
Intron	830. .916	
	/evidence=experimental	
polyA_signal	2177. .2192	
BASE COUNT	651 a 461 c 534 g 570 t	
ORIGIN		
Query Match	89.9%; Score 1834.4; DB 9; Length 2216;	
Best Local Similarity	92.2%; Pred. No. 0;	
Matches 2034; Conservative	0; Mismatches 6; Indels 166; Gaps 3;	
1	CTGCCATTAAGTGGGCGTCCGTGGGGCTGCACAGACAGACTTCCTGTGGGCAGTGGGCTCAG	60
Db	5 CTGCCATTAAGTGGGCGTCCGTGGGGCTGCACAGACAGACTTCCTGTGGGCAGTGGGCTCAG	64
QY	61 GGGCGTCGAACGGAACCCCTACCTCTTGTAGTAGTTCCTGGGGCCACGGCCACGGCAATC	120
Db	65 GGGCGTCGAACGGAACCCCTACCTCTTGTAGTAGTTCCTGGGGCCACGGCCACGGCAATC	124
QY	121 CACGCTGGCGTTGCAGCTTAGGCCACAGCGCTGGCGGCTGAGATCGTCAGCGCTGATCCAT	180
Db	125 CACGCTGGCGTTGCAGCTTAGGCCACAGCGCTGGCGGCTGAGATCGTCAGCGCTGATCCAT	184
QY	181 GCAGGTCATGAAGGCTCAGACATCATATACCAACAAGGTTTCTGGCCAAAGACAGAGAAT	240
Db	185 GCAGGTCATGAAGGCTCAGACATCATATACCAACAAGGTTTCTGGCCAAAGACAGAGAAT	244
QY	241 CTGGCGGCACCAACATGATCAGCTTGTGTGATCCTCTGTGACCAATTACACAGTGTGGA	300
Db	245 CTGGCGGCACCAACATGATCAGCTTGTGTGATCCTCTGTGACCAATTACACAGTGTGGA	304
QY	301 CTTCAAGAAATAGAGCACTGCTCTGTATTTGAGATATATTTGGCCGAGACAAATTCCTAT	360
Db	305 CTTCAAGAAATAGAGCACTGCTCTGTATTTGAGATATATTTGGCCGAGACAAATTCCTAT	364
QY	361 TGTGTGGAGAGAACCAATTTATACATTTGAATCTCTGCTGTGAAAGTCTTGTCAATAC	420
Db	365 TGTGTGGAGAGAACCAATTTATACATTTGAATCTCTGCTGTGAAAGTCTTGTCAATAC	424
QY	421 CAAGCCCCAGAGATGGGCACTGAGAAAAGTGAACCGAAAAGTGGAGCTTGAAGAAAGA	480
Db	425 CAAGCCCCAGAGATGGGCACTGAGAAAAGTGAACCGAAAAGTGGAGCTTGAAGAAAGA	484
QY	481 GGAATGCTTTGTACTTCAAAAAGGCTTAAGCAGGTGGAACCGAAAAGTGGCTGCCAAGT	540
Db	485 GGAATGCTTTGTACTTCAAAAAGGCTTAAGCAGGTGGAACCGAAAAGTGGCTGCCAAGT	544
QY	541 GCATCCACATGCAAAAGCAAAAGTGGCCAGAGGCTTGCAGGTTTGAAGAAACAGGAAT	600
Db	545 GCATCCACATGCAAAAGCAAAAGTGGCCAGAGGCTTGCAGGTTTGAAGAAACAGGAAT	604
QY	601 CTCATCATAGTAATTTCTCCATCGTCAACATACGAGAGAAGTGGTGGCTCCCTTGAGG	660
Db	605 CTCATCATAGTAATTTCTCCATCGTCAACATACGAGAGAAGTGGTGGCTCCCTTGAGG	664

```

QY      2041 A 2041
DB      2041 A 2041

RESULT 2
LOCUS   AX405704                2130 bp    DNA          linear    PAT 14-JUN-2002
DEFINITION Sequence 119 from Patent WO0222660.
ACCESSION AX405704
VERSION  AX405704.1 GI:21438847
KEYWORDS
SOURCE   human.
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
  1. Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,
    Xue, A.J., Yang, Y., Wehrman, F. and Dimanac, R.T.
    Novel nucleic acids and polypeptides
    Patent: WO 0222660-A 119 21-MAR-2002;
    HYSEQ, INC. (US)

FEATURES
    source             1..2130
                        Location/Qualifiers
                        1..2130
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        191..1426
                        /note="unnamed protein product"
                        /codon_start=1
                        /protein_id="CAD34811.1"
                        /db_xref="GI:21438848"
                        /translation="MQVEGLDITNKVSAOERICRHHMISFVDPVNTVTVDFRN
                        RATAIEDIFADKIPVVGNTNYIESILMKVLVTKFOENGTEKVIDRKVELEKED
                        GLVLRKLSQVDPPEMAAKHPDHRKRVASLOVEFTGISHEFLRHROTEGGGPIG
                        GPLKSNPCILWLDADQVLEDRDKRVMDLAAGLELTPHRYNKNKNSENSOD
                        YOHGIFOSIGFEKEFEHFLITEGKCTLETSNOLKGIEMKQVTRAKRKNRYKNR
                        ELSRPPPIPPYVGLSEVSDVSKWSEVLEPALEIYOSFIOGHRPTATPIKMPINEN
                        KRSHYCDLDRILIIIGDREMAHIIKSKHNLKRRRLDSAVNTIESVSYPHNK
                        EPKEKSGPQGNQDELKCSY"

BASE COUNT      620 a      439 c      522 g      549 t
ORIGIN
Query Match      94.1%; Score 1921.2; DB 6; Length 2130;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 2025; Conservative 0; Mismatches 8; Indels 79; Gaps 2;

```

```

DB      380 GGAGAACCAATTATTACATTGATCTCTGCTGGAAATCTCTGCAATACCAAGCC 439
QY      428 CAGAGATGGGCACTGAGAAAGTATGACCGAAAGTGGAGCTTGAAGAGAGATGCT 487
DB      440 CAGAGATGGGCACTGAGAAAGTATGACCGAAAGTGGAGCTTGAAGAGAGATGCT 499
QY      488 CTGTACTTACAAAGCCCTTAAGCCAGGTGGACCCAGAAATGGCTGCCAAGCTGCATCCA 547
DB      500 CTGTACTTACAAAGCCCTTAAGCCAGGTGGACCCAGAAATGGCTGCCAAGCTGCATCCA 559
QY      548 CATGACAAACGCAAGAGTGGCAGAGCTTCAAGTTTGAAGAAACAGCAATCTCAT 607
DB      560 CATGACAAACGCAAGAGTGGCAGAGCTTCAAGTTTGAAGAAACAGCAATCTCAT 619
QY      608 AGTGAATTTCTCATGCTCAACATACGGAAGAGTGGTCCCTTGGAGGCTCTCTG 667
DB      620 AGTGAATTTCTCATGCTCAACATACGGAAGAGTGGTCCCTTGGAGGCTCTCTG 679
QY      668 AAGTCTCTAACCCCTTGATCTCTTGGCTTCATGCTGACCAAGCAGTTTATGATGAGCC 727
DB      680 AAGTCTCTAACCCCTTGATCTCTTGGCTTCATGCTGACCAAGCAGTTTATGATGAGCC 739
QY      728 TTGGATTAAGAGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 787
DB      740 TTGGATTAAGAGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 799
QY      788 CACAGAGCTATATATCAGAAAGTGTTCGAAATATGACGAGACTATCAACATGATATC 847
DB      800 CACAGAGCTATATATCAGAAAGTGTTCGAAATATGACGAGACTATCAACATGATATC 859
QY      848 TTCCATCAATTTGGCTTCAAGAGATTTACAGAGTACTGATCACTGAGGAAATGACACA 907
DB      860 TTCCATCAATTTGGCTTCAAGAGATTTACAGAGTACTGATCACTGAGGAAATGACACA 919
QY      908 CTGAGAGCTATGATACAGCTTCTTAAGAAG----- 938
DB      920 CTGAGAGCTATGATACAGCTTCTTAAGAAGGATTTAGAGCTCTGAACAAAGTACTAAG 939
QY      939 -----GACCTGCTGCC 949
DB      980 AGATATGCCCGGAACAAACCGATGGGTTAAAAACCGTTTTTTGACGACAGCTGGTCCC 1039
QY      950 ATTGTCCCCCTGTCTATGCTTATGAGATATGATGATCTCTGGAAGTGGAGAGTCTGTT 1009
DB      1040 ATTGTCCCCCTGTCTATGCTTATGAGATATGATGATCTCTGGAAGTGGAGAGTCTGTT 1099
QY      1010 CTGGAACCTGCTTGAAGTGTGCAAGTTTATCCAGGCGCAAGAGCTTACAGCCACT 1069
DB      1100 CTGGAACCTGCTTGAAGTGTGCAAGTTTATCCAGGCGCAAGAGCTTACAGCCACT 1159
QY      1070 CCAATAAAGTGCATATCAATGAGAGTGAAGAGAGAGTATATCACTGTGTGACTC 1129
DB      1160 CCAATAAAGTGCATATCAATGAGAGTGAAGAGAGAGTATATCACTGTGTGACTC 1219
QY      1130 TGTGATGGAATCATCATTTGGGATGCGGAATGGGCAAGCAGATTAATCCAAATCCAC 1189
DB      1220 TGTGATGGAATCATCATTTGGGATGCGGAATGGGCAAGCAGATTAATCCAAATCCAC 1279
QY      1190 TTGAACCACTGAAGAAAGAAAGATTTGATGATGATGATGATGATGATGATGATGAT 1249
DB      1280 TTGAACCACTGAAGAAAGAAAGATTTGATGATGATGATGATGATGATGATGATGATGAT 1339
QY      1250 CAGAGTGTTCCTCCAGACTTAAACAAAGAACTAAAGGGAAGGATCCCGAGGCGAGAAAT 1309
DB      1340 CAGAGTGTTCCTCCAGACTTAAACAAAGAACTAAAGGGAAGGATCCCGAGGCGAGAAAT 1399
QY      1310 GATCAAGAGCTGAATATGACAGCTTTAAGAGACATGTCAGAGTGGCTTTGGAAAGTGTG 1369
DB      1400 GATCAAGAGCTGAATATGACAGCTTTAAGAGACATGTCAGAGTGGCTTTGGAAAGTGTG 1459
QY      1370 GGGATCACTTCAAGAGGAGAGGAGTATGTTGATCCAGTCTGGGCAAAAGAGTCTAT 1429
DB      1460 GGGATCACTTCAAGAGGAGAGGAGTATGTTGATCCAGTCTGGGCAAAAGAGTCTAT 1519

```


FEATURES Location/Qualifiers
Source 1..2041
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 589 a 421 c 502 g 529 t
ORIGIN

Query Match 100.0%; Score 2041; DB 6; Length 2041;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCCTAAGATGGCGTCCGTGGCGCTGCACAGACAGTTCCTGTGGGAGTGGCTCG 60
DB 1 CTGCCTAAGATGGCGTCCGTGGCGCTGCACAGACAGTTCCTGTGGGAGTGGCTCG 60

QY 61 GGGCCGTGAAGGACCCCTACCTCTGTAGTGTCTCGGGGGGACAGGGGACGGGAAATC 120
DB 61 GGGCCGTGAAGGACCCCTACCTCTGTAGTGTCTCGGGGGGACAGGGGACGGGAAATC 120

QY 121 CACGCTGGCGTTGCAGCTAGGCCAGCGGCTCGCGGTGAGATCTCAGCGCTGACTCAT 180
DB 121 CACGCTGGCGTTGCAGCTAGGCCAGCGGCTCGCGGTGAGATCTCAGCGCTGACTCAT 180

QY 181 GCAGGTCTATGAAGGCTTAGACATCATCAACAAGGTTTCTGCCAAGACAGAGAAAT 240
DB 181 GCAGGTCTATGAAGGCTTAGACATCATCAACAAGGTTTCTGCCAAGACAGAGAAAT 240

QY 241 CTGGCGGACACATGATAGCTTTGTGGATCTCTGTGACCAATTACAGAGTGTGA 300
DB 241 CTGGCGGACACATGATAGCTTTGTGGATCTCTGTGACCAATTACAGAGTGTGA 300

QY 301 CTTCAAAATAGAGCAACTCTCTGATTAAGATATATTTGCCGAGCAAAATTCCTAT 360
DB 301 CTTCAAAATAGAGCAACTCTCTGATTAAGATATATTTGCCGAGCAAAATTCCTAT 360

QY 361 TTTTGTGGAGAACCAATTATTACATTAATCTCTGCTCTGGAAAGTCTTGTCAATAC 420
DB 361 TTTTGTGGAGAACCAATTATTACATTAATCTCTGCTCTGGAAAGTCTTGTCAATAC 420

QY 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTATTTGACGAAAGTGGAGCTTGAAGAGA 480
DB 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTATTTGACGAAAGTGGAGCTTGAAGAGA 480

QY 481 GGATGGTCTTGTACTTCAAAAAGCCTTAAGCCAGGTGAACCAAGAAATGGTGCACAGT 540
DB 481 GGATGGTCTTGTACTTCAAAAAGCCTTAAGCCAGGTGAACCAAGAAATGGTGCACAGT 540

QY 541 GCATCCACATGACAAAGCGAAAGTGGCCAGAGCTTGCAGAAATTTTGAAGAAACAGGAAT 600
DB 541 GCATCCACATGACAAAGCGAAAGTGGCCAGAGCTTGCAGAAATTTTGAAGAAACAGGAAT 600

QY 601 CTCCTATAGTAATTTCTCATGTCACATACATACGAAAGAGGTGGTCCCTTGGAGG 660
DB 601 CTCCTATAGTAATTTCTCATGTCACATACATACGAAAGAGGTGGTCCCTTGGAGG 660

QY 661 TCCCTGAAGTTCTCTAACCCCTTGACCTTTGGCTTCACTGACAGGACGTTCTAGA 720
DB 661 TCCCTGAAGTTCTCTAACCCCTTGACCTTTGGCTTCACTGACAGGACGTTCTAGA 720

QY 721 TGAAGGCTTGAAGAGGAGGTGATGATGCTGTGCTGGGCTTGGAGAACTAAG 780
DB 721 TGAAGGCTTGAAGAGGAGGTGATGATGCTGTGCTGGGCTTGGAGAACTAAG 780

QY 781 AGATTTTCACAGACGTATTAATCAGAAAGTTTTGGAAATATAGCAGAGACTATCAACA 840
DB 781 AGATTTTCACAGACGTATTAATCAGAAAGTTTTGGAAATATAGCAGAGACTATCAACA 840

QY 841 TGGATCTTCCAAATCAATTGGCTTCAAGAAATTTCAAGATACCTGATCACTGAGGAAA 900
DB 841 TGGATCTTCCAAATCAATTGGCTTCAAGAAATTTCAAGATACCTGATCACTGAGGAAA 900

QY 901 ATGCACATGAGAGACTAGTAACAGCTTCTAAGAAAGAGACTGTCCTCATTTGCCCCC 960
DB 901 ATGCACATGAGAGACTAGTAACAGCTTCTAAGAAAGAGACTGTCCTCATTTGCCCCC 960

DB 901 ATGCACATGAGAGACTAGTAACAGCTTCTAAGAAAGAGACTGTCCTCATTTGCCCCC 960
QY 961 TGTCTATGGCTTAGAGATCTGATGTCTCGAAGTGGAGAGAGTGTCTTGAACCTGC 1020
DB 961 TGTCTATGGCTTAGAGATCTGATGTCTCGAAGTGGAGAGAGTGTCTTGAACCTGC 1020

QY 1021 TCTTGAATTCGTGCAAGTTTCATCCAGGGCCACAAAGCTTACAGCACTCCAAATTAAGAT 1080
DB 1021 TCTTGAATTCGTGCAAGTTTCATCCAGGGCCACAAAGCTTACAGCACTCCAAATTAAGAT 1080

QY 1081 GCCATACATGAGCTGGAACAAAGAAAGTATACCTGTGTGACCTCTGTGATCGAAT 1140
DB 1081 GCCATACATGAGCTGGAACAAAGAAAGTATACCTGTGTGACCTCTGTGATCGAAT 1140

QY 1141 CATCATTTGGGATCCGGAATGGGCGACGCCACATTAATTCCAATTCCTGTAACCAACT 1200
DB 1141 CATCATTTGGGATCCGGAATGGGCGACGCCACATTAATTCCAATTCCTGTAACCAACT 1200

QY 1201 GAAGAAAGAGAGAGATTGGACTGATGCTGTCAACACCATAGAAATCAGAGTGTTC 1260
DB 1201 GAAGAAAGAGAGAGATTGGACTGATGCTGTCAACACCATAGAAATCAGAGTGTTC 1260

QY 1261 CCCAGACTATACAAAGAACCTAAGGGAAGGATCCCGAGGCGCAATGATCAAGAGCT 1320
DB 1261 CCCAGACTATACAAAGAACCTAAGGGAAGGATCCCGAGGCGCAATGATCAAGAGCT 1320

QY 1321 GAAATGAGCGCTTAAAGACATGTCCAGTGGCTTTGGAAGAGTGGGGATCCAGTT 1380
DB 1321 GAAATGAGCGCTTAAAGACATGTCCAGTGGCTTTGGAAGAGTGGGGATCCAGTT 1380

QY 1381 CAGGAGGAGGGGATATGTTGTCTCCAGTCTGGGCAAGAGTCTATGCGGAATCTTC 1440
DB 1381 CAGGAGGAGGGGATATGTTGTCTCCAGTCTGGGCAAGAGTCTATGCGGAATCTTC 1440

QY 1441 TGCATAGCAAGAAAGCTCCACATTTCTTGTATGATGATTTAAAGTCTCAGCTTC 1500
DB 1441 TGCATAGCAAGAAAGCTCCACATTTCTTGTATGATGATTTAAAGTCTCAGCTTC 1500

QY 1501 TATTAATAGAAACAGAGCTTGTGACGCTCTGTGTGCTGATGTCTGAAATGATG 1560
DB 1501 TATTAATAGAAACAGAGCTTGTGACGCTCTGTGTGCTGATGTCTGAAATGATG 1560

QY 1561 TATGTCAGAAAGCATTTTTTTTCTTTGAAACCTTAAAGTCTTATTTAAAGCAGC 1620
DB 1561 TATGTCAGAAAGCATTTTTTTTCTTTGAAACCTTAAAGTCTTATTTAAAGCAGC 1620

QY 1621 ACAGATTCACATTTTATACATGAGAGATCTCTTGTGATGATACAGATGACAGC 1680
DB 1621 ACAGATTCACATTTTATACATGAGAGATCTCTTGTGATGATACAGATGACAGC 1680

QY 1681 ATCCCTTTAAAGAGTTTATGTCCTGACTCTGCTGCTAAATTTCTAATTTCCAGATG 1740
DB 1681 ATCCCTTTAAAGAGTTTATGTCCTGACTCTGCTGCTGCTAATTTCTAATTTCCAGATG 1740

QY 1741 CTTTGTAGATACAGTAATTTTGTGACCCATATTTGGAGTTCTAGATTGAGTGA 1800
DB 1741 CTTTGTAGATACAGTAATTTTGTGACCCATATTTGGAGTTCTAGATTGAGTGA 1800

QY 1801 TGGCAGAAAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAACTAGTTCCGAAT 1860
DB 1801 TGGCAGAAAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAACTAGTTCCGAAT 1860

QY 1861 CTACAGAGAGAGGGAATCAGACTGAGAGAGCTGACATAGGAAGCTTGAAGACCAAGA 1920
DB 1861 CTACAGAGAGAGGGAATCAGACTGAGAGAGCTGACATAGGAAGCTTGAAGACCAAGA 1920

QY 1921 CTTTGAATTTTCCGAGCTCTCATGTGTGATTTATATCATCTGCTCTTTTATTTAGT 1980
DB 1921 CTTTGAATTTTCCGAGCTCTCATGTGTGATTTATATCATCTGCTCTTTTATTTAGT 1980

QY 1981 TACAAATCTAATTTTATTTATGAGTTTAAATGAAGAAATTTTCAAGAAATTTTCAAGAAATTTT 2040
DB 1981 TACAAATCTAATTTTATTTATGAGTTTAAATGAAGAAATTTTCAAGAAATTTTCAAGAAATTTT 2040

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:07:03 ; Search time 6306.89 Seconds

(without alignments)
9418.083 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041

Sequence: 1 CTGCCNTAAGATGCGCTCG.....TTTACAGAGAAAAAAA 2041

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank

1: gb_ba:*
2: gb_ncg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_of:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2041	100.0	2041	6	A98668	A98668 Sequence 3
2	1921.2	94.1	2130	6	AX405704	AX405704 Sequence
3	1834.4	89.9	2216	9	AF074918	AF074918 Homo sapi
4	1556.8	76.3	1749	9	AK000068	AK000068 Homo sapi
5	1556.8	76.3	1749	9	AK074222	AK074222 Homo sapi
6	1489	73.0	1844	9	BC010741	BC010741 Homo sapi
7	1306.4	64.0	1308	9	AY052768	AY052768 Homo sapi
8	1119.4	54.8	2045	10	BC019812	BC019812 Homo sapi
9	850.4	41.7	166496	9	HS118721	HS118721 Human DNA
10	850.4	41.7	200484	2	AC025002	AC025002 Homo sapi
11	839.4	41.1	199863	2	AL354888	AL354888 Homo sapi
12	366	17.9	181838	2	AC106650	AC106650 Homo sapi
13	366	17.9	239888	2	AC093939	AC093939 Rattus no
14	342.6	16.8	355	6	AX336538	AX336538 Sequence
15	273.2	13.4	198772	10	AL606906	AL606906 Mouse DNA
16	249.4	12.2	252	11	G24438	G24438 human STS W
17	224.8	11.0	201861	2	AC103344	AC103344 Rattus no
18	184.8	9.1	199863	2	AL354888	AL354888 Homo sapi
19	160.8	7.9	21571	3	AC014105	AC014105 Drosophila
20	160.8	7.9	21571	3	AC008205	AC008205 Drosophila
21	160.8	7.9	228479	3	AE003749	AE003749 Drosophila
22	131	6.4	77304	2	AC103344	AC103344 Rattus no
23	129.4	6.3	201861	2	AC103344	AC103344 Rattus no
24	126.4	6.2	66680	2	AC100483	AC100483 Mus muscu
25	122.2	6.0	79785	2	AC025068	AC025068 Homo sapi
26	119.8	5.9	77304	2	AC103185	AC103185 Rattus no
27	99	4.9	12242	1	AE013095	AE013095 Thermoplas
28	90.8	4.4	1396	3	AY052773	AY052773 Ctenophora
29	88.2	4.3	993	8	AB062608	AB062608 Arabidops
30	88.2	4.3	1391	8	AB061403	AB061403 Arabidops
31	88.2	4.3	82411	8	AC068809	AC068809 Genomica
32	84.6	4.1	1401	8	AB062609	AB062609 Arabidops
33	84.6	4.1	1632	8	AF109376	AF109376 Arabidops
34	84.6	4.1	1634	8	AY080847	AY080847 Arabidops
35	82.6	4.0	1011	8	AB062610	AB062610 Arabidops
36	82.6	4.0	1067	8	AB061401	AB061401 Arabidops
37	82.6	4.0	1401	8	AY125508	AY125508 Arabidops
38	82.6	4.0	83122	8	ATT20010	ATT20010 Arabidops
39	82.6	4.0	159290	2	AP004813	AP004813 Oryza sat
40	81.2	4.0	38400	1	SC4H2	SC4H2 Streptomy
41	80.8	3.9	161809	2	AP005173	AP005173 Oryza sat
42	78.8	3.9	14486	1	AE004123	AE004123 Vibrio ch
43	78.8	3.9	167350	2	AC129717	AC129717 Oryza sat
44	78.4	3.8	97005	2	AC113584	AC113584 Tetradon
45	78.2	3.8	193068	8	AP003455	AP003455 Oryza sat

ALIGNMENTS

RESULT 1

LOCUS A98668 2041 bp DNA linear PAT 26-JAN-2000

DEFINITION Sequence 3 from Patent WO910482.

ACCESSION A98668

VERSION A98668.1 GI:6781710

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 2041)

AUTHORS Lemieux, J. and Hekimi, S.

TITLE THE C. ELEGANS GRO-1 GENE

JOURNAL Patent: WO 910482-A 3 04-MAR-1999; LEMIEUX JASON (CA); UNIV MCGILL (CA)

Pred. No. is the number of results predicted by chance to have a


```

* 40162 40261: gap of unknown length
* 40262 42177: contig of 1916 bp in length
* 42178 42277: gap of unknown length
* 42278 44344: contig of 2067 bp in length
* 44345 44444: gap of unknown length
* 44445 46472: contig of 2028 bp in length
* 46473 46572: gap of unknown length
* 46573 48054: contig of 1482 bp in length
* 48055 48154: gap of unknown length
* 48155 49390: contig of 1236 bp in length
* 49391 49490: gap of unknown length
* 49491 52986: contig of 3496 bp in length
* 52987 53086: gap of unknown length
* 53087 54647: contig of 1561 bp in length
* 54648 54747: gap of unknown length
* 54748 57112: contig of 2365 bp in length
* 57113 57212: gap of unknown length
* 57213 58448: contig of 1636 bp in length
* 58449 58949: gap of unknown length
* 58949 60817: contig of 1869 bp in length
* 60818 60917: gap of unknown length
* 60918 63443: contig of 2326 bp in length
* 63444 63543: gap of unknown length
* 63544 66294: contig of 2751 bp in length
* 66295 66394: gap of unknown length
* 66395 68140: contig of 1746 bp in length
* 68141 68240: gap of unknown length
* 68241 71198: contig of 2958 bp in length
* 71199 71299: gap of unknown length
* 71299 73990: contig of 2692 bp in length
* 73991 74090: gap of unknown length
* 74091 75872: contig of 1782 bp in length
* 75873 75972: gap of unknown length
* 75973 79070: contig of 3097 bp in length
* 79071 79169: gap of unknown length
* 79170 81462: contig of 2293 bp in length
* 81463 81562: gap of unknown length
* 81563 85047: contig of 3485 bp in length
* 85048 85147: gap of unknown length
* 85148 89223: contig of 4076 bp in length
* 89224 89323: gap of unknown length
* 89324 91525: contig of 2201 bp in length
* 91525 91624: gap of unknown length
* 91625 94768: contig of 3144 bp in length
* 94769 94868: gap of unknown length
* 94869 98303: contig of 3435 bp in length
* 98304 98403: gap of unknown length
* 98404 102480: contig of 4077 bp in length
* 102481 102580: gap of unknown length
* 102581 107197: contig of 4617 bp in length
* 107198 107297: gap of unknown length
* 107298 112447: contig of 5150 bp in length
* 112448 112547: gap of unknown length
* 112548 114968: contig of 2421 bp in length

```

```

Query Match 43.1% Score 38.8; DB 2; Length 161838;
Best Local Similarity 86.0%; Fred. No. 0.0034;
Matches 43; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

Oy 1 TGTGACCTGTGATGATCATCATCTTGGGATGGGATGGGACGCA 50
Db 80342 TGTGACCTTGTGATGATCATCATCTTGGGATGGGACGCA 80391

```

```

RESULT 15
AC093939 239888 bp DNA linear HTG 11-JUL-2002
LOCUS Rattus norvegicus clone CH230-172N14, **** SEQUENCING IN PROGRESS
DEFINITION *** 35 unordered pieces.
AC093939 AC093939.4 GI:21723477
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.

```

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 239888)
Munzy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Aishrooke,S.L., Amaralung,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Bunay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Darhorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabist,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssoen,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lien,C., Liu,D., Liu,W., Louisedge,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maneshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mathney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokwkw,S., Ogutu,M., Okwona,G.,
Oregunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Petersen,L., Pickens,R., Primus,E., Pu,L.L., Qulles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Stutson,A., Swalek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tatney,J., Taylor,C., Taylor,T., Telford,B., Thomas,R., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wellington,S.,
Williams,G., Williamson,A., Wleczek,R., Wooder,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 10, 2002 this sequence version replaced gi:20335375.

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GEXO

Center clone name: CH230-172N14

Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 210507 bases at least Q40

LOCUS AC106650 181838 bp DNA linear HTG 13-JUL-2002
 DEFINITION Rattus norvegicus clone CH230-64124, *** SEQUENCING IN PROGRESS
 *** 62 unordereded pieces.
 AC106650
 AC106650.2 GI:21734730
 HTG: HTGS_PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 181838)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oshan,F.R., Allen,C.,
 Aisbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayale,M., Banks,T.,
 Barberia,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,J.,
 Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhan,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabiell,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hognes,M., Hollway,C., Hollins,B.,
 Homsa,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratoch,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichteige,O., Lieu,C., Liu,J., Liu,W., Louisged,H.,
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhney,E., Mcleod,M.P., Meador,M., Mei,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokwenwo,S., Ogund,M., Okwuonu,G.,
 Ogunyeye,N., Oyiedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,J.,
 Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
 Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,
 Sodergren,E., Sonatke,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Swatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
 Williams,G., Williamson,A., Mleczyk,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G., and Gibbs,R.
 Title Unpublished
 Journal Direct Submission
 Reference 2 (bases 1 to 181838)
 Authors Worley,K.C.
 Journal Title Direct Submission
 Journal Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 181838)
 Reference Direct Submission
 Authors Worley,K.C.
 Journal Title Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 11, 2002 this sequence version replaced gi:18139174.
 Comment
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc.help@bcm.tmc.edu
 ----- Project Information

Center project name: GLTJ
 Center clone name: CH230-64124
 ----- Summary Statistics
 Sequencing vector: M13
 Chemistry: Dye-terminator Big Dye 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 115969 bases at least Q40
 Consensus quality: 122034 bases at least Q30
 Consensus quality: 127422 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 62 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1 1073: contig of 1073 bp in length
 1074 1173: gap of unknown length
 1174 2367: contig of 1194 bp in length
 2368 2467: gap of unknown length
 2468 3755: contig of 1288 bp in length
 3756 3856: gap of unknown length
 3857 5138: contig of 1283 bp in length
 5139 5238: gap of unknown length
 5239 6253: contig of 1016 bp in length
 6254 6354: gap of unknown length
 6355 7644: contig of 1290 bp in length
 7645 7744: gap of unknown length
 7745 9196: contig of 1452 bp in length
 9197 9296: gap of unknown length
 9297 10423: contig of 1127 bp in length
 10424 10523: gap of unknown length
 10524 12093: contig of 1570 bp in length
 12094 12193: gap of unknown length
 12194 13606: contig of 1413 bp in length
 13607 13706: gap of unknown length
 13707 15139: contig of 1433 bp in length
 15140 15239: gap of unknown length
 15240 16519: contig of 1280 bp in length
 16520 16619: gap of unknown length
 16620 18023: contig of 1410 bp in length
 18024 18123: gap of unknown length
 18124 19537: contig of 1408 bp in length
 19538 19637: gap of unknown length
 19638 20773: contig of 1136 bp in length
 20774 20873: gap of unknown length
 20874 21940: contig of 1067 bp in length
 21941 22040: gap of unknown length
 22041 23977: contig of 1937 bp in length
 23978 24077: gap of unknown length
 24078 25464: contig of 1387 bp in length
 25465 25564: gap of unknown length
 25565 26881: contig of 1317 bp in length
 26882 26981: gap of unknown length
 26982 28875: contig of 1894 bp in length
 28876 28975: gap of unknown length
 28976 30545: contig of 1570 bp in length
 30546 30645: gap of unknown length
 30646 31672: contig of 1027 bp in length
 31673 31772: gap of unknown length
 31773 33195: contig of 1423 bp in length
 33196 33295: gap of unknown length
 33296 34682: contig of 1387 bp in length
 34683 34782: gap of unknown length
 34783 35966: contig of 1184 bp in length
 35967 36066: gap of unknown length
 36067 37755: contig of 1689 bp in length
 37756 37855: gap of unknown length
 37856 40161: contig of 2306 bp in length

```
* 28370 28469: gap of 100 bp
* 28470 29153: contig of 684 bp in length
* 28154 29253: gap of 100 bp
* 29254 29971: contig of 718 bp in length
* 29972 30071: gap of 100 bp
* 30072 30770: contig of 699 bp in length
* 30771 30870: gap of 100 bp
* 30871 31581: contig of 711 bp in length
* 31582 31681: gap of 100 bp
* 31682 32380: contig of 699 bp in length
* 32381 32480: gap of 100 bp
* 32481 33187: contig of 707 bp in length
* 33188 33287: gap of 100 bp
* 33288 34012: contig of 725 bp in length
* 34013 34112: gap of 100 bp
* 34113 34831: contig of 719 bp in length
* 34832 34931: gap of 100 bp
* 34932 35627: contig of 696 bp in length
* 35628 35727: gap of 100 bp
* 35728 36446: contig of 719 bp in length
* 36447 36546: gap of 100 bp
* 36547 37245: contig of 699 bp in length
* 37246 37345: gap of 100 bp
* 37346 38077: contig of 732 bp in length
* 38078 38177: gap of 100 bp
* 38178 38879: contig of 702 bp in length
* 38880 38979: gap of 100 bp
* 38980 39688: contig of 709 bp in length
* 39689 39788: gap of 100 bp
* 39789 40500: contig of 712 bp in length
* 40501 40600: gap of 100 bp
* 40601 41317: contig of 717 bp in length
* 41318 41417: gap of 100 bp
* 41418 42116: contig of 699 bp in length
* 42117 42216: gap of 100 bp
* 42217 42923: contig of 707 bp in length
* 42924 43023: gap of 100 bp
* 43024 43739: contig of 716 bp in length
* 43740 43839: gap of 100 bp
* 43840 44570: contig of 731 bp in length
* 44571 44670: gap of 100 bp
* 44671 45381: contig of 711 bp in length
* 45382 45481: gap of 100 bp
* 45482 46191: contig of 710 bp in length
* 46192 46291: gap of 100 bp
* 46292 46986: contig of 695 bp in length
* 46987 47086: gap of 100 bp
* 47087 47789: contig of 703 bp in length
* 47790 47889: gap of 100 bp
* 47890 48590: contig of 701 bp in length
* 48591 48690: gap of 100 bp
* 48691 49414: contig of 724 bp in length
* 49415 49514: gap of 100 bp
* 49515 50220: contig of 706 bp in length
* 50221 50320: gap of 100 bp
* 50321 51010: contig of 690 bp in length
* 51011 51110: gap of 100 bp
* 51111 51823: contig of 713 bp in length
* 51824 51923: gap of 100 bp
* 51924 52648: contig of 725 bp in length
* 52649 52748: gap of 100 bp
* 52749 53466: contig of 718 bp in length
* 53467 53566: gap of 100 bp
* 53567 54295: contig of 729 bp in length
* 54296 54395: gap of 100 bp
* 54396 55105: contig of 710 bp in length
* 55106 55205: gap of 100 bp
* 55206 55935: contig of 730 bp in length
```

```
QY 1 TGTGACCTGTGATGCATTCATTCGGATCGGATGGGACG 46
Db 25567 TGTGACCTGTGATGCATTCATTCGGATCGGATGGGACG 25522
```

```
RESULT 13
AL606906
LOCUS AL606906
DEFINITION Mouse DNA sequence from clone RP23-121J14 on chromosome 4, complete
sequence.
ACCESSION AL606906
VERSION AL606906.18 GI:21955487
KEYWORDS HTG.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 198772)
```

```
REFERENCE Direct Submission
AUTHORS Submitted (25-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
JOURNAL On Jul 25, 2002 this sequence version replaced gi:21912638.
```

COMMENT -----Genome Center

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-121J14 is from the RCT-23 Mouse PAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACe3.6.

FEATURES
source location/Qualifiers
1. 198772

/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-121J14"
/clone_1fb="RCT-23"

BASE COUNT 54889 a 44631 c 45410 g 53842 t
ORIGIN

```
Query Match 44.0%; Score 39.6; DB 10; Length 198772;
Best Local Similarity 91.3%; Pred. No. 0.0017;
Matches 42; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 TGTGACCTGTGATGCATTCATTCGGATCGGATGGGACG 46
|||||
Db 194238 TGTGACCTGTGATGCATTCATTCGGATCGGATGGGACG 194283
```

```
RESULT 14
AC106650
```

* 137461 140476: contig of 3016 bp in length
 * 140477 140576: gap of 100 bp
 Query Match 52.4%; Score 47.2; DB 2; Length 200484;
 Best Local Similarity 86.7%; Pred. No. 2.3e-06;
 Matches 52; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 31 GATCGCATGAGCGACATTAATCCCAATCCAGTGAACCACTGAGAGAGAGA 90
 Db 94041 GATTTTCTTATCAGCGACATTAATCCCAATCCAGTGAACCACTGAGAGAGAGA 93982

RESULT 12
 AC025068/c
 LOCUS
 DEFINITION Homo sapiens chromosome 1 clone RP11-21702 map 1, LOW-PASS SEQUENCE
 SAMPLING.
 AC025068 79785 bp DNA linear HTG 13-JUL-2000
 AC025068 Homo sapiens chromosome 1 clone RP11-21702 map 1, LOW-PASS SEQUENCE
 VERSION AC025068.2 GI:7329438
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Birren, B., Linton, L., Nusbaum, C., and Lander, E.
 TITLE Homo sapiens chromosome 1, clone RP11-21702
 JOURNAL Unpublished
 REFERENCE
 AUTHORS 2 (bases 1 to 79785)
 TITLE Birren, B., Linton, L., Nusbaum, C., and Lander, E.
 JOURNAL Unpublished
 REFERENCE
 AUTHORS 2 (bases 1 to 79785)
 TITLE Birren, B., Linton, L., Nusbaum, C., and Lander, E.
 JOURNAL Unpublished

TITLE
 JOURNAL
 COMMENT
 Submitted (04-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 330 Charles Street, Cambridge, MA 02141, USA
 On Mar 26, 2000 this sequence version replaced gi:7158126.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIRB
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L7716
 Center clone name: 217_O_2

* NOTE: This record contains 98 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.

* However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved. 740: contig of 740 bp in length
 1
 741 840: gap of 100 bp
 841 1543: contig of 703 bp in length
 1544 1643: gap of 100 bp
 1644 2350: contig of 707 bp in length
 2351 2450: gap of 100 bp
 2451 3158: contig of 708 bp in length
 3159 3258: gap of 100 bp
 3259 3978: contig of 720 bp in length
 3979 4078: gap of 100 bp
 4079 4775: contig of 698 bp in length
 4777 4876: gap of 100 bp
 4877 5593: contig of 717 bp in length
 5594 5693: gap of 100 bp
 5694 6412: contig of 719 bp in length
 6413 6512: gap of 100 bp
 6513 7238: contig of 726 bp in length
 7239 7338: gap of 100 bp
 7339 8055: contig of 717 bp in length
 8056 8155: gap of 100 bp
 8156 8870: contig of 715 bp in length
 8871 8970: gap of 100 bp
 8971 9691: contig of 721 bp in length
 9692 9791: gap of 100 bp
 9792 10493: contig of 702 bp in length
 10494 10593: gap of 100 bp
 10594 11326: contig of 733 bp in length
 11327 11426: gap of 100 bp
 11427 12137: contig of 711 bp in length
 12138 12237: gap of 100 bp
 12238 12950: contig of 713 bp in length
 12951 13050: gap of 100 bp
 13051 13751: contig of 701 bp in length
 13752 13851: gap of 100 bp
 13852 14552: contig of 701 bp in length
 14553 14652: gap of 100 bp
 14653 15372: contig of 720 bp in length
 15373 15472: gap of 100 bp
 15473 16183: contig of 711 bp in length
 16184 16283: gap of 100 bp
 16284 17005: contig of 722 bp in length
 17006 17105: gap of 100 bp
 17106 17825: contig of 720 bp in length
 17826 17925: gap of 100 bp
 17926 18639: contig of 714 bp in length
 18640 18739: gap of 100 bp
 18740 19457: contig of 718 bp in length
 19458 19557: gap of 100 bp
 19558 20268: contig of 712 bp in length
 20270 20369: gap of 100 bp
 20370 21069: contig of 700 bp in length
 21070 21169: gap of 100 bp
 21170 21895: contig of 726 bp in length
 21896 21995: gap of 100 bp
 21996 22708: contig of 714 bp in length
 22710 22809: gap of 100 bp
 22810 23507: contig of 698 bp in length
 23508 23607: gap of 100 bp
 23608 24297: contig of 690 bp in length
 24298 24397: gap of 100 bp
 24398 25107: contig of 710 bp in length
 25108 25207: gap of 100 bp
 25208 25936: contig of 729 bp in length
 25937 26036: gap of 100 bp
 26037 26749: contig of 713 bp in length
 26750 26849: gap of 100 bp
 26850 27562: contig of 713 bp in length
 27563 27662: gap of 100 bp
 27663 28369: contig of 707 bp in length

* NOTE: This is a 'working draft' sequence. It currently
* consists of 81 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved. 1048: contig of 1048 bp in length
* 1
* 1049 1148: gap of 100 bp
* 1149 2470: contig of 1322 bp in length
* 2471 2570: gap of 100 bp
* 2571 3643: contig of 1073 bp in length
* 3644 3743: gap of 100 bp
* 3744 4980: contig of 1237 bp in length
* 4981 5080: gap of 100 bp
* 5081 6352: contig of 1272 bp in length
* 6353 6452: gap of 100 bp
* 6453 7549: contig of 1097 bp in length
* 7550 7649: gap of 100 bp
* 7650 8977: contig of 1328 bp in length
* 8978 9077: gap of 100 bp
* 9078 10103: contig of 1026 bp in length
* 10104 10203: gap of 100 bp
* 10204 11362: contig of 1159 bp in length
* 11363 11462: gap of 100 bp
* 11463 12821: contig of 1359 bp in length
* 12822 12921: gap of 100 bp
* 12922 14007: contig of 1086 bp in length
* 14008 14107: gap of 100 bp
* 14108 15283: contig of 1176 bp in length
* 15284 15383: gap of 100 bp
* 15384 16399: contig of 1016 bp in length
* 16400 16499: gap of 100 bp
* 16500 17804: contig of 1305 bp in length
* 17805 17904: gap of 100 bp
* 17905 19176: contig of 1272 bp in length
* 19177 19276: gap of 100 bp
* 19277 20384: contig of 1108 bp in length
* 20385 20484: gap of 100 bp
* 20485 21689: contig of 1205 bp in length
* 21690 21789: gap of 100 bp
* 21790 23191: contig of 1402 bp in length
* 23192 23291: gap of 100 bp
* 23292 24513: contig of 1222 bp in length
* 24514 24613: gap of 100 bp
* 24614 26135: contig of 1522 bp in length
* 26136 26235: gap of 100 bp
* 26236 27373: contig of 1138 bp in length
* 27374 27473: gap of 100 bp
* 27474 28723: contig of 1250 bp in length
* 28724 28823: gap of 100 bp
* 28824 30217: contig of 1394 bp in length
* 30218 30317: gap of 100 bp
* 30318 32056: contig of 1739 bp in length
* 32057 32156: gap of 100 bp
* 32157 33818: contig of 1662 bp in length
* 33819 33918: gap of 100 bp
* 33919 35156: contig of 1238 bp in length
* 35157 35256: gap of 100 bp
* 35257 36712: contig of 1456 bp in length
* 36713 36812: gap of 100 bp
* 36813 38379: contig of 1567 bp in length
* 38380 38479: gap of 100 bp
* 38480 40190: contig of 1711 bp in length
* 40191 40290: gap of 100 bp
* 40291 41780: contig of 1490 bp in length
* 41781 41880: gap of 100 bp
* 41881 44053: contig of 2173 bp in length
* 44054 44153: gap of 100 bp
* 44154 45474: contig of 1321 bp in length
* 45475 45574: gap of 100 bp
* 45575 46821: contig of 1247 bp in length

* 46822 46921: gap of 100 bp
* 46922 47302: contig of 381 bp in length
* 47303 47402: gap of 100 bp
* 47403 49323: contig of 1921 bp in length
* 49324 49423: gap of 100 bp
* 49424 51140: contig of 1717 bp in length
* 51141 51240: gap of 100 bp
* 51241 53099: contig of 1859 bp in length
* 53100 53199: gap of 100 bp
* 53200 54631: contig of 1432 bp in length
* 54632 54731: gap of 100 bp
* 54732 55751: contig of 1020 bp in length
* 55752 55851: gap of 100 bp
* 55852 57470: contig of 1619 bp in length
* 57471 57570: gap of 100 bp
* 57571 60182: contig of 2612 bp in length
* 60183 60282: gap of 100 bp
* 60283 63065: contig of 2783 bp in length
* 63066 63165: gap of 100 bp
* 63166 64443: contig of 1278 bp in length
* 64444 64543: gap of 100 bp
* 64544 66043: contig of 1500 bp in length
* 66044 66143: gap of 100 bp
* 66144 68623: contig of 2480 bp in length
* 68624 68723: gap of 100 bp
* 68724 70122: contig of 1399 bp in length
* 70123 70222: gap of 100 bp
* 70223 72407: contig of 2185 bp in length
* 72408 72507: gap of 100 bp
* 72508 74635: contig of 2128 bp in length
* 74636 74735: gap of 100 bp
* 74736 77554: contig of 2819 bp in length
* 77555 77654: gap of 100 bp
* 77655 80277: contig of 2623 bp in length
* 80278 80377: gap of 100 bp
* 80378 81798: contig of 1421 bp in length
* 81799 81898: gap of 100 bp
* 81899 84126: contig of 2228 bp in length
* 84127 84226: gap of 100 bp
* 84227 86794: contig of 2568 bp in length
* 86795 86894: gap of 100 bp
* 86895 90311: contig of 3417 bp in length
* 90312 90411: gap of 100 bp
* 90412 92255: contig of 1844 bp in length
* 92256 92355: gap of 100 bp
* 92356 94570: contig of 2215 bp in length
* 94571 94670: gap of 100 bp
* 94671 97163: contig of 2493 bp in length
* 97164 97263: gap of 100 bp
* 97264 99724: contig of 2461 bp in length
* 99725 99824: gap of 100 bp
* 99825 103151: contig of 3327 bp in length
* 103152 103251: gap of 100 bp
* 103252 106370: contig of 3119 bp in length
* 106371 106470: gap of 100 bp
* 106471 108888: contig of 2418 bp in length
* 108889 108988: gap of 100 bp
* 108989 111514: contig of 2526 bp in length
* 111515 111614: gap of 100 bp
* 111615 114360: contig of 2746 bp in length
* 114361 114460: gap of 100 bp
* 114461 118269: contig of 3809 bp in length
* 118270 118369: gap of 100 bp
* 118370 122359: contig of 3990 bp in length
* 122360 122459: gap of 100 bp
* 122460 125669: contig of 3210 bp in length
* 125670 125769: gap of 100 bp
* 125770 129306: contig of 3537 bp in length
* 129307 129406: gap of 100 bp
* 129407 132984: contig of 3578 bp in length
* 132985 133084: gap of 100 bp
* 133085 137360: contig of 4276 bp in length
* 137361 137460: gap of 100 bp

FEATURES
source

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

31 GATCGGATGGCGACACATAAATCCAATCCACTTGACCCTGAAGAAAGA 90

Db 26154 GATTTGTCTTATCAGCGCACATAAAATCCAATCCCACTTGAACCACTGAAGAAAGA 26095

RESULT 11

	Homo sapiens	chromosome 1 clone RP11-204L3 map 1,	*** SEQUENCING INFORMATION ***
DEFINITION			

AC025002.3	GT:8516086
ACCESSION	AC025002
VERSION	

SOURCE	HOMO sapiens.
ORGANISM	HOMO sapiens

REFERENCE
1 (bases 1 to 300484)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

TITLE	ISSN	NUMBER	DATE	BY	FOR	REMARKS
Homo sapiens chromosome 1, clone RP11-204L3						
troubled						

AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,

Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,

Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczky, J.,

Meldrum, J.: Meneus, B.: Mihova, T.: Miranda, C.: Mjenqa, V.: Morrow, J.:
Mucciarini, M.: Mucchan, F.: Mucchar, A.: Muckenham, N.: Muenkelers, N.:

O'Neill, D.; Oliva, I.M.; Oliver, C.; Petersoll, K.; Pierre, N.; Pisaní, C.; Pollara, V.; Raymond, C.; Riley, B.; Rogov, P.; Rothman, D.

Testave, S., Theodore, J., A., Travers, M., Trigg, J.,
Stange-Johnson, N., Stojanovic, N., Sudriamallahi, A.,
Tadamas, J.,

TITLE
Young, G., Zalnoun, J., Zimmer, A. and Zody, M.
Direct Submission

On Tue 14 2000 this secure worder replaced it: 7770400
Research, 320 Charles Street, Cambridge, MA 02141, USA

Smit, A.F.A. & Green, P. (1996-1997)

Center: Whitehead Institute/ MIT Center for Genome Research

Contact: sequence_submissions@genome.wi.mit.edu

Center clone name: 204_L_3

```

/evidence-not_experimental
repeat_region 12351..13402 /note="AluX repeat: matches 19..300 of consensus"
/note="L2 repeat: matches 957..2141 of consensus"
repeat_region 13474..13591 /note="L2 repeat: matches 1468..1747 of consensus"
/note="AluX/AluY repeat: matches 178..295 of consensus"
repeat_region 13596..14185 /note="L2 repeat: matches 2244..2278 of consensus"
/note="L1M3 repeat: matches 7071..7691 of consensus"
repeat_region 14186..14425 /note="L2 repeat: matches 2129..2417 of consensus"
/note="L2 repeat: matches 27..147 of consensus"
repeat_region 14457..14572 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 47..137 of consensus"
repeat_region 15137..15215 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 16285..16510 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 17313..17597 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 20320..20782 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 20785..20932 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 21140..21425 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 21927..22073 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 22185..22408 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 22409..22704 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 22705..22759 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 22829..22898 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 22932..22989 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 22990..23298 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 23299..23381 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 23401..23576 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 23600..23742 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 24157..24451 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 24618..24684 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 24753..24891 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 25018..25159 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 26215..26275 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 26604..26643 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 26930..27021 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 27380..27733 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 28059..28151 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
misc_feature 28542..30030 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
/evidence-not_experimental
repeat_region 29451..29584 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 30126..30262 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 30283..30460 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 30928..31015 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 31389..31662 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"

Query Match 52.4%, Score 47.2, DB 9, Length 166496;
Best Local Similarity 86.7%, Pred. No. 2.2e-06;
Matches 52; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 31 GATCGGATGGGCGGACATTAATCCACTGACCACTGAAGAAAGA 90
DB 82459 GATTGTGCTTATGACGACATTAATCCACTGACCACTGAAGAAAGA 82400

RESULT 10
AL354888/c 199863 bp DNA linear HTG 13-JUN-2001
LOCUS Homo sapiens chromosome 1 clone RP11-473P22, ** SEQUENCING IN
DEFINITION PROGRESS **, 16 unordered pieces.
ACCESSION AL354888
VERSION AL354888.6 GI:9863692
KEYWORDS HTG: HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 199863)
REFERENCE
AUTHORS Plumb, B.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Aug 21, 2000 this sequence version replaced gi:9213379.
COMMENT ----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA473P22
----- Summary Statistics
Assembly program: XGAP4, version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 189476 bases at least Q40
Consensus quality: 193311 bases at least Q30
Consensus quality: 195510 bases at least Q20
Insert size: 198363; sum-of-contigs
Insert size: 201159; 6.2% error; agarose-fp
Quality coverage: 3.87x in Q20 bases; sum-of-contigs Quality
coverage: 3.98x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2278: contig of 2278 bp in length
* 2279 2378: gap of 100 bp
* 2379 22743: contig of 20365 bp in length
* 22744 22843: gap of 100 bp

```


TITLE Novel nucleic acids and polypeptides
JOURNAL Patent: WO 022660-A 119 21-MAR-2002;
HYSEO, INC. (US)

FEATURES
source
Location/Qualifiers
1..2130
/organism="Homo sapiens"
/db_xref="taxon:9606"
191..1426
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD34811.1"
/db_xref="gi:21438848"
/translation="MQVEGLDIITNKVSAQEOIRCHHMSFVDPVLYNTVYDFERN
RATLIEDIFARDKPIVGGTYLIESLKMVYVTKPOEMGETEKYIDRVELEKED
GLVHKRLSVDPEMAAKLHPHDKRKVARSIOVEPFGISHSFELHROHREGGGPIG
GPLKESNPCLIMLHADQAVLDERLDRVDDMLAGLIEELRDFHRRINQKAVSENSOD
YHGIFOSIGFEFEHYLITEGKCTLETSNQLKKEIALKQVTKRKARKONVKNR
FLSRGPVIVPVYGLESDVSKWESVLEPALLEIVOSFIQGHKPTATPIKMPYNAENR
KRSYHLCDCRRIIGDREMAAHIKSKSHLNQKKRRRLSDAVNTIESQSVSDHNK
EPKESPGONDQELKCSV"

BASE COUNT 620 a 439 c 522 g 549 t

ORIGIN

Query Match 100.0%; Score 90; DB 6; Length 2130;
Best Local Similarity 100.0%; Pred. No. 8.2e-23;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCAATGGCGACGACATAAAATCC 60
|||||
Db 1211 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCAATGGCGACGACATAAAATCC 1270
|||||

OY 61 AAATCCACTTGAACCACTGAAGAAAAGA 90
|||||
Db 1271 AAATCCACTTGAACCACTGAAGAAAAGA 1300
|||||

RESULT 7
AF074918 2216 bp RNA linear PRI 13-DEC-2000
LOCUS Homo sapiens tRNA isopentenylpyrophosphate transferase precursor
DEFINITION
ACCESSION AF074918
VERSION AF074918.1 GI:11337964
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Golovko, A., Hjalms, G., Stibon, F. and Nilander, B.
TITLE Cloning of a human tRNA isopentenyl transferase
JOURNAL Gene 258 (1-2), 85-93 (2000)
MEDLINE 20564178
PUBMED 11111046
REFERENCE 2 (bases 1 to 2216)
AUTHORS Golovko, A. and Hjalms, G.
TITLE Direct Submission
JOURNAL Submitted (28-JUN-1998) Dep. of Plant Biology, Swedish University
of Agricultural Sciences, Box 7080, SE-750 07 Uppsala, Sweden
Location/Qualifiers
1..2216
/organism="Homo sapiens"
/db_xref="dbEST:AA332152"
/db_xref="taxon:9606"
/note="extension of sequence by 5' RACE"
1..2216
/note="incompletely processed mRNA"
join(15..829,917..1505)
/note="liphase"
/codon_start=1
/product="tRNA isopentenylpyrophosphate transferase"
/protein_id="AAG31324.1"
/db_xref="GI:11337965"

/translation="MASVAARAAPVSGSLRGILQRTLPVILGATGTCKSTLALQGL
ORLGEIVASDSQVYEGDIIITNKVSAQEOIRCHHMSFVDPVLYNTVYDFERN
TALIEDIFARDKPIVGGTYLIESLKMVYVTKPOEMGETEKYIDRVELEKEDGL
VLHKRLSVDPEMAAKLHPHDKRKVARSIOVEPFGISHSFELHROHREGGGPIG
LKSNPCLIMLHADQAVLDERLDRVDDMLAGLIEELRDFHRRINQKAVSENSODY
HGIFOSIGFEFEHYLITEGKCTLETSNQLKKEIALKQVTKRKARKONVKNR
FLSRGPVIVPVYGLESDVSKWESVLEPALLEIVOSFIQGHKPTATPIKMPYNAENR
SYHLCDCRRIIGDREMAAHIKSKSHLNQKKRRRLSDAVNTIESQSVSDHNK
EPKESPGONDQELKCSV"

BASE COUNT 651 a 461 c 534 g 570 t

ORIGIN

Query Match 100.0%; Score 90; DB 9; Length 2216;
Best Local Similarity 100.0%; Pred. No. 8.3e-23;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCAATGGCGACGACATAAAATCC 60
|||||
Db 1290 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCAATGGCGACGACATAAAATCC 1349
|||||

OY 61 AAATCCACTTGAACCACTGAAGAAAAGA 90
|||||
Db 1350 AAATCCACTTGAACCACTGAAGAAAAGA 1379
|||||

RESULT 8
BC019812 2045 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, RIKEN cDNA 2310075G14 gene, clone MGC:30541
DEFINITION IMAGE:5042856, mRNA, complete cds.
ACCESSION BC019812
VERSION BC019812.1 GI:18044185
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: gcgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarsson, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,
Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAP Plate: 41 Row: 1 Column: 20.
Location/Qualifiers
1..2045
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:30541 IMAGE:5042856"
/tissue_type="Kidney, normal, 5 month old male mouse."
/clone_id="NCI_CGAP_Kid14"

```

source
1. 1749
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL01371"
/tissue_type="colon"
/clone_lib="COL"
/notes="cloning vector: pME18SFL3"

BASE COUNT      541 a      337 c      408 g      463 t
ORIGIN

Query Match      100.0%; Score 90; DB 9; Length 1749;
Best Local Similarity 100.0%; Pred. No. 8e-23;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGAGCGCATATAATCC 60
      |||||||
DB 826 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGAGCGCATATAATCC 885
      |||||||

QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
      |||||||
DB 886 AATCCCACTTGAACCACTGAAGAAAAGA 915
      |||||||

RESULT 4
BC010741      1844 bp      mRNA      linear      PRI 17-JUL-2001
LOCUS      Homo sapiens, Similar to tRNA isopentenylpyrophosphate transferase,
ACCESSION      BC010741
VERSION      BC010741.1 GI:14789610
KEYWORDS      MGC.
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1844)
AUTHORS      Strausberg, R.
TITLE      Direct Submision
JOURNAL      Submitted (12-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAK Plate: 14 Row: P Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA 91: 7019914.
Location/Qualifiers
1. 1844
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:17002 IMAGE:3905836"
/tissue_type="uterus, leiomyosarcoma"
/clone_lib="NIH_MGC_71"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
166. 1140
/codon_start=1
/product="Similar to tRNA isopentenylpyrophosphate
transferase"
CDS

```

```

/protein_id="AAH10741.1"
/db_xref="GI:14789611"
/translation="MGTKEVIDRKVELKEEDGIVLHKRLSDVPEMAAKLHPDRKY
ARSLGVFEETGSHSEFLRQHTREEGGPIKSPKSNFCILMLHADQDEBLDRVD
DMLAGLIEELDFSHFRYRNOKVNSNDQYQGIQSIGKFEHEYLITREGCITETS
NOLKRGITALKOVTRKTRAKONRWKNFSLRPGPIVYPVIGLEYSVSKNEESTYLE
PALEIVQSFIOGHKPTATPIKMPYNAENKRKSYHLCDLDRILIGREVAHIKSKSH
LNOLKRRRLRDSDDAVNTTIESQSVSPDHNEPREKESPGNDDELKCSV"

BASE COUNT      545 a      374 c      454 g      471 t
ORIGIN

Query Match      100.0%; Score 90; DB 9; Length 1844;
Best Local Similarity 100.0%; Pred. No. 8e-23;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGAGCGCATATAATCC 60
      |||||||
DB 925 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGAGCGCATATAATCC 984
      |||||||

QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
      |||||||
DB 985 AATCCCACTTGAACCACTGAAGAAAAGA 1014
      |||||||

RESULT 5
A98668      2041 bp      DNA      linear      PAT 26-JAN-2000
LOCUS      Sequence 3 from Patent WO9910482.
DEFINITION      A98668
ACCESSION      A98668
VERSION      A98668.1 GI:6781710
KEYWORDS      unidentified.
SOURCE      unidentified.
ORGANISM      unidentified.

REFERENCE      1 (bases 1 to 2041)
AUTHORS      Lemieux, J., and Hekimi, S.
TITLE      THE C. ELEGANS GRO-1 GENE
JOURNAL      Patent: WO 9910482-A 3 04-MAR-1999;
LEMIEUX JASON (CA); UNIV MCGILL (CA)
Location/Qualifiers
FEATURES
source
1. 2041
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT      589 a      421 c      502 g      529 t
ORIGIN

Query Match      100.0%; Score 90; DB 6; Length 2041;
Best Local Similarity 100.0%; Pred. No. 8.2e-23;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGAGCGCATATAATCC 60
      |||||||
DB 1121 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGAGCGCATATAATCC 1180
      |||||||

QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
      |||||||
DB 1181 AATCCCACTTGAACCACTGAAGAAAAGA 1210
      |||||||

RESULT 6
AX405704      2130 bp      DNA      linear      PAT 14-JUN-2002
LOCUS      Sequence 119 from Patent WO0222260.
DEFINITION      AX405704
ACCESSION      AX405704
VERSION      AX405704.1 GI:21438847
KEYWORDS      human.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F.,
Xue, A. J., Yang, Y., Wehrman, T. and Drmanac, R. T.

```

trna-modifying enzyme in the mitochondria
Genetics 159 (1), 147-157 (2001)

JOURNAL
MEDLINE
21444833
PUBMED
11560893

REFERENCE
2 (bases 1 to 1308)
AUTHORS
Lemieux, J., Barnes, T. and Hekimi, S.
TITLE
Direct Submission
JOURNAL
Submitted (28-AUG-2001) Biology, McGill University, 1205 Dr.
Penfield Avenue, Montreal, QC H3A 1B1, Canada

FEATURES
source
1..1308
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-2ec05"
/note="H901"
/codon_start=1
/product="tRNA isopentenyl transferase"
/protein_id="AA114107.1"
/db_xref="gi:16209579"
/translation="AAVAVGSGELGRLPTLVVILGATGTGKSTLALDGLGRLGE
IVSADSMQVYEGLDITNKVSAQEOIRCHMISFVDPVNTNVTVDPRNATLIED
IFARDKIPVVGNTNYIESLIMKVLVNTKQEMETEKVIDRKYLEKEDGIVLKRLL
SOVPEMAAKLPHDKRKVARSLOVEETGISHSEFLHROHTEEGGGLPKRFSNP
CIWLHADOAVLDERLDRVDMGLAAGLLELRDRHRYNOKNNSENSODVOHGIFOS
IGFKEFHEYLITEGKCTLETSNOLKKGPGPIVPPVIGLEVSADVSKWESVLEPLET
VOSTIQGKRPTATPIKMPYNAENKRSYHLCDLDRILIGREMAAHIKSKSHLMOLK
RRRLDSQAVNTISOSVSPDHNEKPEKESPGQNDLCKSV"

BASE COUNT
ORIGIN
378 a 302 c 329 g 299 t

Query Match
Best Local Similarity 100.0%; Score 90; DB 9; Length 1308;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGATCGATCATTCATTTGGGATCGGAATGGGACGACATAAATCC 60
|||||
DB 1093 TGTGACCTCTGATCGATCATTCATTTGGGATCGGAATGGGACGACATAAATCC 1152
|||||

QY 61 AAATCCACTTGACCACTGAAGAAAGA 90
|||||
DB 1153 AAATCCACTTGACCACTGAAGAAAGA 1182
|||||

RESULT 2
LOCUS AK000068 1749 bp mRNA linear PRI 22-FEB-2000
DEFINITION Homo sapiens cDNA FLJ20061 fls, clone COL01383.
ACCESSION AK000068
VERSION AK000068.1 GI:7019914
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens colon cDNA to mRNA, clone_1lb.COL clone:COL01383.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites)
AUTHORS Kawabata, A., Hiki, T., Kobatake, N., Inagaki, H., Ikema, Y.,
Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,
Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1749)
AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T.,
Shibahara, T., Tanaka, T. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@elms.u-tokyo.ac.jp,
Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of

Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (Partly supported by Science and Technology
Agency).

FEATURES
source
1..1749
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL01383"
/tissue="colon"
/clone_1lb="COL"
/note="cloning vector pME18SFL3"
61..1041
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA90923.1"
/db_xref="gi:7019915"
/translation="MGTEKVIDRKYLEKEDGIVLHRLSOVPEMAKIPPHKRY
ARSLQVEPETGISHSEFLHROHTEEGGGLPKRFSNPCLIMLHADOAVLDERLDR
VDMLAAGLLELRDRHRYNOKNNSENSODYQHGIFOSIGFKEFHEYLITEGKCTLE
TSNOLKRGIALKQVTRVARKONRWKNNRSLRPGPIVPPVIGLEVSADVSKWESV
LEPALEIVOSTIQGKRPTATPIKMPYNAENKRSYHLCDLDRILIGREMAAHIKSK
SHLMOLKRRRLDSQAVNTISOSVSPDHNEKPEKESPGQNDLCKSV"

BASE COUNT
ORIGIN
541 a 337 c 408 g 463 t

Query Match
Best Local Similarity 100.0%; Score 90; DB 9; Length 1749;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGATCGATCATTCATTTGGGATCGGAATGGGACGACATAAATCC 60
|||||
DB 826 TGTGACCTCTGATCGATCATTCATTTGGGATCGGAATGGGACGACATAAATCC 885
|||||

QY 61 AAATCCACTTGACCACTGAAGAAAGA 90
|||||
DB 886 AAATCCACTTGACCACTGAAGAAAGA 915
|||||

RESULT 3
LOCUS AK074222 1749 bp mRNA linear PRI 15-FEB-2002
DEFINITION Homo sapiens cDNA FLJ2642 fls, clone COL01371, highly similar to
Homo sapiens tRNA isopentenylpyrophosphate transferase precursor
RNA.
ACCESSION AK074222
VERSION AK074222.1 GI:18676765
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens colon cDNA to mRNA, clone_1lb.COL clone:COL01371.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites)
AUTHORS Kawabata, A., Hiki, T., Kobatake, N., Inagaki, H., Ikema, Y.,
Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,
Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1749)
AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T.,
Shibahara, T., Tanaka, T. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@elms.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan: cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction,
5'- & 3'-end one pass sequencing: Department of Virology and Human
Genome Center, Institute of Medical Science, University of Tokyo
(Partly supported by Science and Technology Agency).

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:07:03 ; Search time 278.109 Seconds

(without alignments)
9418.083 Million cell updates/sec

Title: US-09-513-151-3_COPY_1121_1210

Sequence: 1 TGTGACCTCTGTGATCGAAT.....TGACCACTGAGAAAGA 90

Scoring table: IDENTIFY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sta:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_mu:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sta:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	1308	9 AY052768	AY052768 Homo sapi
2	90	100.0	1749	9 AK000068	AK000068 Homo sapi
3	90	100.0	1749	9 AK074222	AK074222 Homo sapi
4	90	100.0	1844	9 BC010741	BC010741 Homo sapi
5	90	100.0	2041	6 A98668	A98668 Sequence 3
6	90	100.0	2130	6 AX405704	AX405704 Sequence
7	90	100.0	2216	9 AF074918	AF074918 Homo sapi
8	74	82.2	2045	10 BC019812	BC019812 Mus muscu
9	47.2	52.4	166496	9 HS118171	AL035272 Human DNA
10	47.2	52.4	199863	2 AL354888	AL354888 Homo sapi
11	47.2	52.4	200484	2 AC025002	AC025002 Homo sapi
12	46	51.1	79785	2 AC025068	AC025068 Homo sapi
13	39.6	44.0	198772	10 AL606906	AL606906 Mouse DNA
14	38.8	43.1	181838	2 AC106650	AC106650 Rattus no
15	38.8	43.1	239888	2 AC093939	AC093939 Rattus no
16	38	42.2	201861	2 AC103344	AC103344 Rattus no
17	31.4	34.9	149452	2 AC124912	AC124912 Papio cyn
18	30.8	34.2	61294	2 AC107873	AC107873 Homo sapi
19	30.8	34.2	120007	9 AF064864	AF064864 Homo sapi
20	30.8	34.2	340000	9 HS21C083	AL163283 Homo sapi
21	30.4	33.8	278652	2 AC073809	AC073809 Mus muscu
22	29.4	32.7	118444	2 AC114896	AC114896 Oryza sat
23	29.4	32.7	320250	9 AF117829	AF117829 Homo sapi
24	29	32.2	55227	2 AC100331	AC100331 Mus muscu
25	29	32.2	55227	2 AC100331	AC100331 Mus muscu
26	29	32.2	66821	2 AC121146	AC121146 Mus muscu
27	29	32.2	175390	2 AC114551	AC114551 Mus muscu
28	29	32.2	229957	10 AL670231	AL670231 Mouse DNA
29	28.8	32.0	163115	9 AC105902	AC105902 Homo sapi
30	28.8	32.0	200956	9 AC104165	AC104165 Homo sapi
31	28.6	31.8	152505	2 AC121658	AC121658 Rattus no
32	28.6	31.8	186793	2 AC095657	AC095657 Rattus no
33	28.6	31.8	198839	2 AC094648	AC094648 Rattus no
34	28.4	31.6	167439	2 AC130273	AC130273 Papio cyn
35	28.4	31.6	176562	2 AC130272	AC130272 Papio cyn
36	28.4	31.6	190352	2 AC101735	AC101735 Mus muscu
37	28.4	31.6	210563	2 AC117723	AC117723 Mus muscu
38	28.2	31.3	41633	9 AC096585	AC096585 Homo sapi
39	28.2	31.3	74531	2 AC015829	AC015829 Homo sapi
40	28.2	31.3	97255	9 AL590482	AL590482 Human DNA
41	28.2	31.3	114041	2 AP001866	AP001866 Homo sapi
42	28.2	31.3	137091	2 AC025484	AC025484 Homo sapi
43	28.2	31.3	162289	4 AC097230	AC097230 Sus scrofa
44	28.2	31.3	163460	2 AC069378	AC069378 Homo sapi
45	28.2	31.3	166039	2 AC044855	AC044855 Homo sapi

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
AY052768	AY052768	Homo sapiens tRNA isopentenyl transferase mRNA, partial cds.	AY052768	1	GI:16209578	Homo sapiens.	Homo sapiens.	1308 bp mRNA linear PRI 17-OCT-2001	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	Regulation of physiological rates in Caenorhabditis elegans by a

Pred. No. is the number of results predicted by chance to have a

Db 28 GAGGCTCAGAGAGCTTATAATCATGGCAGAGAGCGAAGGGAAGCAAGACCTTCTTCA 87

Qy 62 AATCCCACTTGAACCACTGAAGAAAAGA 90

Db 88 CATGGCAGCAGGAGAGAAAGAGAAAGGA 116

Search completed: April 21, 2003, 13:37:15
Job time : 30.58 secs


```

OY      8  TCTGTGATGCAATGATCTTGGGATCGCGAATGGCGCACATTAATTCGAATCC 67
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      1647 TTTTATTTCACACATCTCTTTAGGTGTGACGTCGTGTAAAGCAAAATTAATCTTAAGCC 1706
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY      68  ACTTGACCACTGAAGAAA 87
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      1707 CCCCCAACCATCTGTAATGGA 1726

RESULT 10
ID      AAT71699 standard; DNA; 20303 BP.
XX      AAT71699:
DT      20-AUG-1997 (first entry)
DE      Human deoxycytidylate deaminase Intron 2 encoding DNA.
XX      Recombinant deaminase; dCMP; ds.
XX      Homo sapiens.
XX      OS
XX      US5622851-A.
XX      PD
XX      22-APR-1997.
XX      PE
XX      10-JAN-1995; 95US-0370975.
XX      PR
XX      10-JAN-1995; 95US-0370975.
XX      PA
XX      (HEAL-) HEALTH RES INC.
XX      PI
XX      Maley F, Maley GR, Weiner KYB;
XX      WPI: 1997-244391/22.
XX      DR
XX      DNA encoding human deoxycytidylate deaminase - for production of
XX      recombinant deaminase
XX      PS
XX      Claim 2; Column 83-100; 58pp; English.
XX      CC
XX      The present sequence encodes the human deoxycytidylate (dCMP)
XX      deaminase Intron 2, which comprises 20303 base pairs from nucleotides
XX      1964-22266 of the dCMP deaminase sense strand. The dCMP deaminase gene
XX      contains a 5' untranslated region (including the promoter), 5 exons,
XX      4 introns and a 3' untranslated region (including the stop signals).
XX      The gene can be used to produce recombinant dCMP deaminase, which can
XX      be used to convert dCMP to dUMP. Also, the dCMP gene can be altered
XX      (removed or mutated) to alter DNA replication in cells, which may lead
XX      to mutagenesis.
XX      CC
XX      Sequence 20303 BP; 5454 A; 4115 C; 5052 G; 5682 T; 0 other;
XX      SO

Query Match      29.8%; Score 26.8; DB 18; Length 20303;
Best Local Similarity 57.0%; Pred. No. 15;
Matches 49; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

OY      3  TGACCTCTGTGATCGAATCATTCATTGGGATCGCGAATGGCGCACATTAATTCGA 62
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      4437 TGACATTTTTCATTCCTTTTCTTGTTGTGAATTCATTATAGTGAAGCACATTTAATCAAA 4486
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

OY      63  ATCCCACTGGAACCAATGAAGAAA 88
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      4487 ATACCAATGTTATCAAAATGAGGTAAA 4512

RESULT 11
ID      AAT71696
XX      AAT71696 standard; DNA; 26764 BP.
XX      AAT71696:
AC

```

XX	20-AUG-1997	(first entry)
DT		
XX		
DE	Human deoxycytidylate deaminase gene.	
XX		
KW	Recombinant deaminase; dCMP; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	misc-feature	1..1317
FT	exon	/tag= a
FT		/note= "5' untranslated region, including promotor"
FT		1318..1425
FT	Intron	/tag= b
FT		/number= 1
FT		1426..1827
FT		/tag= c
FT	exon	/number= 1
FT		1828..1963
FT		/tag= d
FT	Intron	/number= 2
FT		1964..2226
FT		/tag= e
FT	exon	/number= 2
FT		22267..22383
FT		/tag= f
FT	Intron	/number= 3
FT		22384..23740
FT		/tag= g
FT	exon	/number= 3
FT		23741..23837
FT		/tag= h
FT		/number= 4
FT	Intron	23838..25391
FT		/tag= i
FT	exon	/number= 4
FT		25392..25467
FT		/tag= j
FT		/number= 5
FT	misc-feature	25468..26764
FT		/tag= k
FT		/note= "3' untranslated region"
XX		
PN	US5622851-A.	
XX		
PD	22-APR-1997.	
XX		
PE	10-JAN-1995;	95US-0370975.
XX		
PR	10-JAN-1995;	95US-0370975.
XX		
PA	(HEAL-) HEALTH RES INC.	
XX		
PI	Maley F, Maley GR, Weiner KXB;	
XX		
DR	WP1; 1997-244391/22.	
DR	P-PSDB; AAW18205.	
XX		
PT	DNA encoding human deoxycytidylate deaminase - for production of recombinant deaminase	
XX		
PS	Claim 3; Column 55-78; 58pp; English.	
XX		
CC	The present sequence encodes the human deoxycytidylate (dCMP) deaminase gene which contains a 5' untranslated region (including the promoter), 5 exons, 4 introns and a 3' untranslated region (including the stop signals). The gene can be used to produce recombinant dCMP deaminase, which can be used to convert dCMP to dUMP. Also, the dCMP gene can be altered (removed or mutated) to alter DNA replication in cells, which may lead to mutagenesis.	
CC		
Sequence	26764 BP; 7079 A; 5521 C; 6539 G; 7625 T; 0 other;	

PR 14-AUG-2000: 2000US-0225447.
 PR 14-AUG-2000: 2000US-0225757.
 PR 14-AUG-2000: 2000US-0225758.
 PR 14-AUG-2000: 2000US-0225759.
 PR 18-AUG-2000: 2000US-0226279.
 PR 22-AUG-2000: 2000US-0226681.
 PR 22-AUG-2000: 2000US-0226682.
 PR 23-AUG-2000: 2000US-0227182.
 PR 30-AUG-2000: 2000US-0227009.
 PR 30-AUG-2000: 2000US-0228924.
 PR 01-SEP-2000: 2000US-0229287.
 PR 01-SEP-2000: 2000US-0229343.
 PR 01-SEP-2000: 2000US-0229344.
 PR 01-SEP-2000: 2000US-0229345.
 PR 05-SEP-2000: 2000US-0229513.
 PR 05-SEP-2000: 2000US-0229519.
 PR 06-SEP-2000: 2000US-0230437.
 PR 06-SEP-2000: 2000US-0230438.
 PR 08-SEP-2000: 2000US-0231242.
 PR 08-SEP-2000: 2000US-0231243.
 PR 08-SEP-2000: 2000US-0231244.
 PR 08-SEP-2000: 2000US-0231245.
 PR 08-SEP-2000: 2000US-0231413.
 PR 08-SEP-2000: 2000US-0231414.
 PR 08-SEP-2000: 2000US-0232080.
 PR 12-SEP-2000: 2000US-0232081.
 PR 14-SEP-2000: 2000US-0232397.
 PR 14-SEP-2000: 2000US-0232398.
 PR 14-SEP-2000: 2000US-0232399.
 PR 14-SEP-2000: 2000US-0232400.
 PR 14-SEP-2000: 2000US-0232401.
 PR 14-SEP-2000: 2000US-0233063.
 PR 14-SEP-2000: 2000US-0233064.
 PR 14-SEP-2000: 2000US-0233065.
 PR 21-SEP-2000: 2000US-0234223.
 PR 21-SEP-2000: 2000US-0234274.
 PR 25-SEP-2000: 2000US-0234997.
 PR 25-SEP-2000: 2000US-0234998.
 PR 25-SEP-2000: 2000US-0235484.
 PR 27-SEP-2000: 2000US-0235834.
 PR 27-SEP-2000: 2000US-0235836.
 PR 27-SEP-2000: 2000US-0236327.
 PR 29-SEP-2000: 2000US-0236367.
 PR 29-SEP-2000: 2000US-0236368.
 PR 29-SEP-2000: 2000US-0236369.
 PR 29-SEP-2000: 2000US-0236370.
 PR 02-OCT-2000: 2000US-0236802.
 PR 02-OCT-2000: 2000US-0237037.
 PR 02-OCT-2000: 2000US-0237038.
 PR 02-OCT-2000: 2000US-0237039.
 PR 02-OCT-2000: 2000US-0237040.
 PR 13-OCT-2000: 2000US-0239935.
 PR 13-OCT-2000: 2000US-0239937.
 PR 20-OCT-2000: 2000US-0240960.
 PR 20-OCT-2000: 2000US-0241121.
 PR 20-OCT-2000: 2000US-0241785.
 PR 20-OCT-2000: 2000US-0241786.
 PR 20-OCT-2000: 2000US-0241787.
 PR 20-OCT-2000: 2000US-0241808.
 PR 20-OCT-2000: 2000US-0241809.
 PR 20-OCT-2000: 2000US-0241826.
 PR 01-NOV-2000: 2000US-0244617.
 PR 08-NOV-2000: 2000US-0246474.
 PR 08-NOV-2000: 2000US-0246475.
 PR 08-NOV-2000: 2000US-0246476.
 PR 08-NOV-2000: 2000US-0246477.
 PR 08-NOV-2000: 2000US-0246523.
 PR 08-NOV-2000: 2000US-0246524.
 PR 08-NOV-2000: 2000US-0246525.
 PR 08-NOV-2000: 2000US-0246526.
 PR 08-NOV-2000: 2000US-0246527.
 PR 08-NOV-2000: 2000US-0246528.

PR 08-NOV-2000: 2000US-0246532.
 PR 08-NOV-2000: 2000US-0246609.
 PR 08-NOV-2000: 2000US-0246610.
 PR 08-NOV-2000: 2000US-0246611.
 PR 08-NOV-2000: 2000US-0246613.
 PR 17-NOV-2000: 2000US-0249207.
 PR 17-NOV-2000: 2000US-0249208.
 PR 17-NOV-2000: 2000US-0249209.
 PR 17-NOV-2000: 2000US-0249210.
 PR 17-NOV-2000: 2000US-0249211.
 PR 17-NOV-2000: 2000US-0249212.
 PR 17-NOV-2000: 2000US-0249213.
 PR 17-NOV-2000: 2000US-0249214.
 PR 17-NOV-2000: 2000US-0249215.
 PR 17-NOV-2000: 2000US-0249216.
 PR 17-NOV-2000: 2000US-0249217.
 PR 17-NOV-2000: 2000US-0249218.
 PR 17-NOV-2000: 2000US-0249244.
 PR 17-NOV-2000: 2000US-0249245.
 PR 17-NOV-2000: 2000US-0249264.
 PR 17-NOV-2000: 2000US-0249265.
 PR 17-NOV-2000: 2000US-0249297.
 PR 17-NOV-2000: 2000US-0249299.
 PR 17-NOV-2000: 2000US-0249300.
 PR 01-DEC-2000: 2000US-0250160.
 PR 01-DEC-2000: 2000US-0250391.
 PR 05-DEC-2000: 2000US-0251030.
 PR 05-DEC-2000: 2000US-0251988.
 PR 05-DEC-2000: 2000US-0256719.
 PR 06-DEC-2000: 2000US-0251479.
 PR 08-DEC-2000: 2000US-0251856.
 PR 08-DEC-2000: 2000US-0251868.
 PR 08-DEC-2000: 2000US-0251869.
 PR 08-DEC-2000: 2000US-0251989.
 PR 08-DEC-2000: 2000US-0251990.
 PR 11-DEC-2000: 2000US-0254097.
 PR 05-JAN-2001: 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI: 2001-483426/52.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis.

XX Disclosure; SEQ ID NO 31146; 3071pp + Sequence Listing; English.

XX AA54951 to AA64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AA54951 to AA64702. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AA64703
 CC to AA87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AA54942 to AA54950 and AA87699
 CC represent sequences used in the exemplification of the present invention.

XX Sequence 3616 BP; 1000 A; 855 C; 784 G; 977 T; 0 other;

Query Match 30.2%; Score 27.2; DB 22; Length 3616;
 Best Local Similarity 58.8%; Pred. No. 6.7; Indels 0; Gaps 0;
 Matches 47; Conservative 0; Mismatches 33;

CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC	sequences (AB101840-AB116175) and the encoded proteins
CC	(AB57737-AB572072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
XX	Sequence 6399 BP: 1596 A; 1565 C; 1565 G; 1673 T; 0 other;
XX	
XX	Query Match
XX	Best Local Similarity 30.4%; Score 27.4; DB 23; Length 6399;
XX	Matches 43; Conservative 62.3%; Pred. No. 6.6;
XX	Mismatches 26; Indels 0; Gaps 0;
QY	1 TGTGACCTCTGTGATGCATCATCTATTGGGATCGCGAATGGCGACGACATAAATCC 60
DB	3921 TGTCAAAATATGCGAAGCGCATTTCTGTGGGAGTACCAATGGGAGCTCATGTGAAGTCC 3862
QY	61 AAATCCGAC 69
DB	3861 AACCAACAC 3853
XX	
XX	RESULT 9
XX	AAK76334
XX	AAK76334 standard; DNA; 3616 BP.
XX	
XX	AAK76334;
XX	
XX	07-NOV-2001 (first entry)
XX	
XX	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:31146.
XX	
XX	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX	cytostatic; gene therapy; vaccine; metastasis; ds.
XX	
XX	Homo sapiens.
XX	
XX	WO200157182-A2.
XX	
XX	09-AUG-2001.
XX	
XX	17-JAN-2001; 2001WO-US01354.
XX	
XX	31-JAN-2000; 2000US-0179065.
XX	04-FEB-2000; 2000US-0180628.
XX	24-FEB-2000; 2000US-0184664.
XX	02-MAR-2000; 2000US-0186350.
XX	16-MAR-2000; 2000US-0189874.
XX	17-MAR-2000; 2000US-0190076.
XX	18-APR-2000; 2000US-0198123.
XX	19-MAY-2000; 2000US-0205515.
XX	07-JUN-2000; 2000US-0209467.
XX	28-JUN-2000; 2000US-0214886.
XX	30-JUN-2000; 2000US-0215135.
XX	07-JUL-2000; 2000US-0216647.
XX	11-JUL-2000; 2000US-0216880.
XX	11-JUL-2000; 2000US-0217487.
XX	11-JUL-2000; 2000US-0217496.
XX	14-JUL-2000; 2000US-0218290.
XX	26-JUL-2000; 2000US-0220963.
XX	26-JUL-2000; 2000US-0220964.
XX	14-AUG-2000; 2000US-0224518.
XX	14-AUG-2000; 2000US-0224519.
XX	14-AUG-2000; 2000US-0225213.
XX	14-AUG-2000; 2000US-0225214.
XX	14-AUG-2000; 2000US-0225266.
XX	14-AUG-2000; 2000US-0225267.
XX	14-AUG-2000; 2000US-0225268.
XX	14-AUG-2000; 2000US-0225270.

(II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 1776 BP; 491 A; 450 C; 464 G; 371 T; 0 other;

Query Match 30.7%; Score 27.6; DB 23; Length 1776;
Best Local Similarity 60.8%; Pred. No. 3.9;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCAGCGACATAAAATCC 60
|||||
Db 117 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCAGCGACATAAAATCC 176

QY 61 AATATCCCACTTGAA 74
| | | | |
Db 177 TCAACACACTGCAA 190

RESULT 6

AAS67229 standard; cDNA; 1780 BP.

XX AAS67229;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #3033.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HSE-) HXSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR P-PSDB: ABG03042.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 1; SEQ ID No 3033; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 1780 BP; 491 A; 452 C; 465 G; 372 T; 0 other;

Query Match 30.7%; Score 27.6; DB 23; Length 1780;
Best Local Similarity 60.8%; Pred. No. 3.9;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCAGCGACATAAAATCC 60
|||||
Db 117 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCAGCGACATAAAATCC 176

QY 61 AATATCCCACTTGAA 74
| | | | |
Db 177 TCAACACACTGCAA 190

RESULT 7

AAS55928 standard; DNA; 2703 BP.

XX AAS55928;

DT 13-FEB-2002 (first entry)

DE Streptococcus pneumoniae DNA for cellular proliferation protein #499.

KW Antisense; ds; prokaryotic cellular proliferation gene;

KW antibiotic; antibacterial; drug design.

OS Streptococcus pneumoniae.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207272P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;
DR WPI: 2001-611495/70.
DR P-PSDB: AAU038069.

XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
PS Claim 27; Seq ID No 9565; 511pp; English.

CC genes correlated with a cancerous state of a mammalian cell, comprising
 CC detecting at least one differentially expressed gene product in a test
 CC sample derived from a cell suspected of being cancerous, where detection
 CC of the differentially expressed gene product is correlated with a
 CC cancerous state of the cell from which the test sample was derived.
 CC The polynucleotide sequences can be used in a method for detecting
 CC differentially expressed genes correlated with a cancerous state of a
 CC mammalian cell. The polynucleotides can also be used as probes for
 CC detecting and mapping related genes. They can be used in diagnosis and
 CC prognosis of diseases and disorders (e.g. identification of
 CC pre-metastatic or metastatic cancerous states, stages of cancer, or
 CC responsiveness of cancer to therapy). This is particularly for breast
 CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
 CC negative breast cancer, lung cancer, and colon cancer.

XX Sequence 300 BP; 83 A; 69 C; 72 G; 76 T; 0 other;

Query Match 48.9%; Score 44; DB 21; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.4e-06;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTTGGGGATCGGAATGGC 44
 DB 257 TGTGACCTCTGTGATCGAATCATTTGGGGATCGGAATGGC 300

RESULT 4
 AAS67227
 ID AAS67227 standard; cDNA; 1248 BP.

AC AAS67227;

DT 13-FEB-2002 (first entry).

DE DNA encoding novel human diagnostic protein #3031.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR P-PSDB; ABG03040.

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity

PS Claim 1; SEQ ID NO 3031; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1248 BP; 334 A; 343 C; 289 G; 282 T; 0 other;

Query Match 30.7%; Score 27.6; DB 23; Length 1248;
 Best Local Similarity 60.8%; Pred. No. 3.6;

Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTTGGGGATCGGAATGGCAGCGCATATAATCC 60
 DB 423 TGTGACCTCTGTGTACAAAGCAGGATATAGGGAAGGAGAAATATCTCC 482

OY 61 AATCCGACTTGA 74
 DB 483 TCAACACACTGCAA 496

RESULT 5
 AAS64323
 ID AAS64323 standard; cDNA; 1776 BP.

AC AAS64323;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #127.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR P-PSDB; ABG00136.

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity

PS Claim 1; SEQ ID NO 127; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving

PT useful for enhancing longevity of a host and inhibiting tumour
PT formation
XX
XX Claim 8; Fig 8; 93pp; English.
XX
XX The present sequence represents the human homologue of gro-1, and
CC is referred to as hgro-1. The specification describes the five genes
CC of the Caenorhabditis elegans gro-1 operon (AA36071). The operon
CC contains the gro-1 gene (AA36072), the gop-1 gene (AA36074), the gop-2
CC gene (AA36075), the gop-2 gene (AA36075), and the hap-1 gene
CC (AA36077). The gro-1 gene can be used in a method for the diagnosis
CC and/or prognosis of cancer in a patient. Transgenic mice containing a
CC gene knock-out of a murine gene homologue of the gro-1 gene are useful
CC as models of aging and cancer. The proteins encoded by the genes are
CC useful for identifying compounds that affect the enzymatic activity
CC of these proteins. In order to enhance longevity of a host and inhibit
CC tumour formation, the gro-1 gene, together with the gop-1, gop-2,
CC gop-3 and hap-1 genes enables study of a physiological clock.
SQ Sequence 2041 BP; 589 A; 421 C; 502 G; 529 T; 0 other;
QY
Query Match 100.0%; Score 90; DB 20; Length 2041;
Best Local Similarity 100.0%; Pred. No. 6e-23;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1121 TGTGACCTCTGTGATCGAATCATCATTCATGGGATCGCGAATGGCGACACATTAATCC 60
QY 1 TGTGACCTCTGTGATCGAATCATCATTCATGGGATCGCGAATGGCGACACATTAATCC 60
DB 1121 TGTGACCTCTGTGATCGAATCATCATTCATGGGATCGCGAATGGCGACACATTAATCC 1180
QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
DB 1181 AATCCCACTTGAACCACTGAAGAAAAGA 1210
RESULT 2
ABNS9708
ID ABNS9708 standard; cDNA; 2130 BP.
XX
AC ABNS9708;
XX
XX 28-JUN-2002 (first entry)
DE
XX Novel human coding sequence SEQ ID NO: 119.
XX
XX Human: anti-naemic; vulnerary; anti-inflammatory; immunomodulator;
KM anti-infectivity; cerebroprotective; cytoskeletal; rheumatic; gene therapy;
KM neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200222660-A2.
XX
XX 21-MAR-2002.
XX
XX 10-SEP-2001; 2001WO-US26015.
XX
XX 11-SEP-2000; 2000US-0659671.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F,
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX WPI; 2002-292408/33.
DR P-PSDB; ABB97295.
XX
XX An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis -
XX
XX Claim 1; SEQ ID NO 119; 509pp; English.
XX
XX The present invention provides the protein and coding sequences of 444

CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate actin or inhibit e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention.
SQ Sequence 2130 BP; 620 A; 439 C; 522 G; 549 T; 0 other;
QY
Query Match 100.0%; Score 90; DB 24; Length 2130;
Best Local Similarity 100.0%; Pred. No. 6.1e-23;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1211 TGTGACCTCTGTGATCGAATCATCATTCATGGGATCGCGAATGGCGACACATTAATCC 1270
QY 1 TGTGACCTCTGTGATCGAATCATCATTCATGGGATCGCGAATGGCGACACATTAATCC 60
DB 1211 TGTGACCTCTGTGATCGAATCATCATTCATGGGATCGCGAATGGCGACACATTAATCC 1270
QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
DB 1271 AATCCCACTTGAACCACTGAAGAAAAGA 1300
RESULT 3
AAA00911
ID AAA00911 standard; cDNA; 300 BP.
XX
AC AAA00911;
XX
XX 19-MAY-2000 (first entry)
DE
XX Human colon cancer cell line polynucleotide sequence SEQ ID NO:902.
XX
XX Human: colon cancer; tumour; diagnosis; gene expression product;
KM probe; detection; cancerous state; metastasis; identification;
KM breast cancer; oestrogen receptor-positive breast cancer; therapy;
KM oestrogen receptor-negative breast cancer; lung cancer; ss.
XX
XX Homo sapiens.
XX
XX WO958675-A2.
XX
XX 18-NOV-1999.
XX
XX 13-MAY-1999; 99WO-US10602.
XX
XX 14-MAY-1998; 98US-0085426.
XX
XX 15-MAY-1998; 98US-0085537.
XX
XX 15-MAY-1998; 98US-0085696.
XX
XX 21-OCT-1998; 98US-0105234.
XX
XX 27-OCT-1998; 98US-0105877.
XX
XX (CHIR) CHIRON CORP.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Glese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lamson G, Drmanac R, Crtvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kila D, Garcia V, Jones LM, Stache-Crain B;
XX
XX WPI; 2000-126369/11.
DR
XX Polynucleotide library used to determine cancerous states of mammalian
PT cells -
XX
XX Claim 1; Page 399; 1097pp; English.
XX
XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
CC libraries constructed from human colon cancer cell lines. The present
CC invention also describes a method of detecting differentially expressed

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:03:18 ; Search time 24.58 Seconds

(without alignments)
8245.719 Million cell updates/sec

Title: US-09-513-151-3_COPY_1121_1210

Perfect score: 90
Sequence: 1 TCGAGACCTCTGTGATCGCAT.....TGACCACTGAAGAAAGA 90

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002:*

1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	2041	20	AAAX36073 Human homologue of
2	90	100.0	2130	24	ABNS9708 Human colon cancer
3	44	48.9	300	21	AAA00911 Human encoding novel
4	27.6	30.7	1248	23	AA567227 DNA encoding novel
5	27.6	30.7	1776	23	AA564323 DNA encoding novel
6	27.6	30.7	1780	23	AA567229 DNA encoding novel
7	27.4	30.4	2703	23	AA555928 Streptococcus pneu
8	27.4	30.4	6399	23	ABL17248 Drosophila melanog
9	27.2	30.2	3616	22	AAK76334 Human immune/haema

10	26.8	29.8	20303	18	AAT71699 Human deoxyctidyl
11	26.8	29.8	26764	18	AAT71696 Human deoxyctidyl
12	26.6	29.6	255	19	AAAX11458 Human baillietic po
13	26.6	29.6	255	19	AAAX11457 Human baillietic po
14	26.6	29.6	344	20	AAH87247 Human single nucle
15	26.6	29.6	344	20	AAH87261 Human single nucle
16	26.2	29.1	3364	23	ABL05458 Drosophila melanog
17	26.2	29.1	5299	23	ABL07890 Drosophila melanog
18	26	28.9	1472	23	AA575217 DNA encoding novel
19	26	28.9	2640	22	AA529615 Human endocrine po
20	26	28.9	3423	22	AA158249 Human polynucleoti
21	26	28.9	3443	20	AA508694 Novel nucleotide s
22	26	28.9	3747	23	ABL12801 Drosophila melanog
23	26	28.9	4346	22	AA56756 Human cardiovascu
24	26	28.9	11116	19	AA159270 Human polynucleoti
25	25.8	28.7	861	19	AAV43021 Streptococcus pneu
26	25.8	28.7	2793	21	AA609487 Streptococcus pneu
27	25.4	28.2	1815	20	AA560607 Polynucleotide seq
28	25.4	28.2	1815	20	AA560608 Polynucleotide seq
29	25.4	28.2	1815	22	AA500566 Truncated cryf to
30	25.4	28.2	1815	22	AA500567 Truncated cryf to
31	25.4	28.2	7028	23	ABL30122 Drosophila melanog
32	25.4	28.2	11313	23	ABL12356 Drosophila melanog
33	25.4	28.2	20776	23	ABL30040 Drosophila melanog
34	25.2	28.0	90	24	ABR36689 Human DNA encoding
35	25.2	28.0	408	23	AA588099 Human DNA encoding
36	25.2	28.0	471	20	AA532049 Human MERT2 relate
37	25.2	28.0	471	22	AA590306 HEONN73R cDNA clon
38	25.2	28.0	502	24	AA38928 Exon 1 of S1g1ec8-
39	25.2	28.0	626	21	AA43007 Human 5' EST ISOLA
40	25.2	28.0	1090	22	AA494448 Human foetal cDNA,
41	25.2	28.0	1554	24	ABR43364 DNA encoding stail
42	25.2	28.0	1592	24	AA38935 S1g1ec8-L coding c
43	25.2	28.0	1597	20	AA517773 SAR-2 polypeptide
44	25.2	28.0	2898	22	AA522575 Human cDNA encodin
45	25.2	28.0	2898	22	AA522811 Human cDNA encodin

ALIGNMENTS

RESULT 1
ID AAX36073 standard; DNA: 2041 BP.
XX AAX36073:
AC
XX
XX
DT 15-JUL-1999 (first entry)
XX
DE Human homologue of gro-1, referred to as hgro-1.
XX
XX gro-1 operon; gro-1 gene; gop-1 gene; gop-2 gene;
KW hap-1 gene; cancer; aging; longevity; tumour formation;
KW physiological clock; ss.
XX
XX Homo sapiens.
OS
XX
XX WO910482-A1.
XX
XX
PD 04-MAR-1999.
XX
XX 20-AUG-1998; 98WO-CA00803.
XX
XX 25-AUG-1997; 97CA-2210251.
XX
XX
PA (UYMC-) UNIV MCGILL.
XX
XX Barnes T, Hekimi S, Lakowski B, Lemieux J;
PI WPI; 1999-190615/16.
XX
XX
XX Molecular identity of the gro-1 gene - useful for cancer diagnosis
PT and/or prognosis, and where compounds affecting encoded proteins are

```
; FEATURE:      Description:      Comment:
```

Search completed: April 21, 2003, 16:30:12
Job time : 12.3078 secs


```

      : FILING DATE:
      : ATTORNEY/AGENT INFORMATION:
      : NAME: Brookes, A. Anders
      : REGISTRATION NUMBER: 36,373
      : REFERENCE/DOCKET NUMBER: PB340P1
      : TELECOMMUNICATION INFORMATION:
      : TELEPHONE: (301) 309-8504
      : TELEFAX: (301) 309-8512
      : INFORMATION FOR SEQ ID NO: 210:
      : SEQUENCE CHARACTERISTICS:
      : LENGTH: 11378 base pairs
      : TYPE: nucleic acid
      : STRANDEDNESS: double
      : TOPOLOGY: linear
      : US-08-961-527-210
    Query Match          26.9%; Score 24.2; DB 4; Length 11378;
    Best Local Similarity 55.3%; Pred. No. 12;
    Matches 47; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
  QY   2 GTGACCTGTGCATCGAATCATCATTTGGGATGCCGAATGGCCAGCCACATAAATCCA 61
       ||| -||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     6393 GTGTATCGTGATGTCGTGTCGTTACTATCATATGATGATGTAAGTCCAGCTTGAACA 6452
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY     62 AATCCACTGTGACCACACTGAAGA 86
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     6453 AAGACGACGTGAAGAAGTAGACAA 6477

RESULT 12
US-08-484-891-7/C
Sequence 7, Application US/08484891
Patent No. 5935935
GENERAL INFORMATION:
APPLICANT: Connolly, Sheila
APPLICANT: Kaleko, Michael
APPLICANT: Smith, Theodore
TITLE OF INVENTION: Adenoviral Vectors for
TITLE OF INVENTION: Treatment of Hemophilia
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bahn, Gilfillan,
ADDRESSEE: Cecchi, Stewart & Olstein
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: New Jersey
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,891
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/218,335
FILING DATE: 25-MAR-1994
APPLICATION NUMBER: 08/074,920
FILING DATE: 10-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Olstein, Elliot M.
REGISTRATION NUMBER: 24,025
REFERENCE/DOCKET NUMBER: 271010-273
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4629 bases
TYPE: nucleic acid

```

```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA primer
FEATURE:
NAME/KEY: Factor VII cDNA with
NAME/KEY: B domain deleted
US-08-484-891-7

Query Match      26.7%; Score 24; DB 2; Length 4629;
Best Local Similarity 58.3%; Pred. No. 10;
Matches 42; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY   18 AATCATCTTGGGATCGGCAGATGGGCGACGCACTAAATCCCAATTGGAACA 77
Db    2446 AATCTTCCTTTCTTACTTTCACACTGATGATGATCATTCACTAGTAATTCCTTTATCTG 2387
              ||||| | | | | | | | | | | | | | | | | | | | | | | |
QY     78 ACTGAAGAAG 89
Db    2386 ACTGAAGTAG 2375
               ||||| | |

RESULT 13
US-08-717-294-41/c
Sequence 41, Application US/08717294
Patent No. 6114148
GENERAL INFORMATION:
APPLICANT: SEED, BRIAN
APPLICANT: HAAS, JURGEN
TITLE OF INVENTION: HIGH LEVEL EXPRESSION OF
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 110
CORRESPONDENCE ADDRESSES:
ADDRESS: Clark & Ebling LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/717,294
FILING DATE: 20-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ebling, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/345001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 4670 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-717-294-41

Query Match      26.7%; Score 24; DB 3; Length 4670;
Best Local Similarity 58.3%; Pred. No. 10;
Matches 42; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY   18 AATCATCTTGGGATCGGCAGATGGGCGACGCACTAAATCCCAATTGGAACA 77

```

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA
; APPLICATION NUMBER:

```

```

1 FILING DATE:
2
3 CLASSIFICATION: 424
4
5 PRIOR APPLICATION DATA:
6
7 APPLICATION NUMBER:
8

```

```

: GENERAL INFORMATION:
: APPLICANT: Zalacain, Magdalena
: APPLICANT: Burnham, Martin K. R.
: APPLICANT: Biswas, Sanjoy
: APPLICANT: Brown, James R.
: APPLICANT: Ingraham, Karen A.
: APPLICANT: Chalker, Allison F.
: APPLICANT: So, Chl Young
: APPLICANT: Holmes, David J.
: APPLICANT: Trainl, Christopher M.
: APPLICANT: Warren, Richard L.
: APPLICANT: Mathie, Thomas B.
: TITLE OF INVENTION: infb
: FILE REFERENCE: GM10175
: CURRENT APPLICATION NUMBER: US/09/574,912
: CURRENT FILING DATE: 2000-05-19
: PRIOR APPLICATION NUMBER: 09/283,763
: PRIOR FILING DATE: 1999-04-01
: PRIOR APPLICATION NUMBER: 60/105,985
: PRIOR FILING DATE: 1999-10-28
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 2793
: TYPE: DNA
: ORGANISM: Streptococcus pneumoniae
: US-09-574-912-1

```

[illegible]

```

RESULT 6
US-09-178-252-7/C
; Sequence 7, Application US/09178252
; Patent No. 6218188
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Stelman, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XC2
; CURRENT APPLICATION NUMBER: US/09/178,252
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 60/065,215
; EARLIER FILING DATE: 1997-11-12
; EARLIER APPLICATION NUMBER: 60/076,445
; EARLIER FILING DATE: 1998-03-02
; NUMBER OF SEQ. ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1815
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic B.t. toxin gene
; US-09-178-252-7

```

Query Match	28.2%	Score 25.4	DB 4	Length 1815
Best Local Similarity	60.6%	Pred. No. 2.1		
Matches 35; Conservative	0;	Mismatches 16;	Indels 0;	Gaps 0;
3 TGACCTCTGATCGAATCATCATTTGGGGATCGCAATGGGAGCGACAT	53			

```

Db      430  TGATGGCTGTGATCAAGCATCATCTGTGTAGCAAGCGTATGGCAGCAT 380

      RESULT 7
      US-09-178-252-8/c
      : Sequence 8, Application US/09178252
      : Patent No. 6218188
      :
      : GENERAL INFORMATION:
      : APPLICANT: Cardineau, Guy A.
      : APPLICANT: Steلمان, Steven J.
      : APPLICANT: Narva, Kenneth E.
      : TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
      : FILE REFERENCE: MA-714XC2
      : CURRENT APPLICATION NUMBER: US/09/178,252
      : CURRENT FILING DATE: 1998-10-23
      : EARLIER APPLICATION NUMBER: 60/065,215
      : EARLIER FILING DATE: 1997-11-12
      : EARLIER APPLICATION NUMBER: 60/076,445
      : EARLIER FILING DATE: 1998-03-02
      : NUMBER OF SEQ ID NOS: 27
      : SOFTWARE: PatentIn Ver. 2.0
      : SEQ ID NO 8
      :
      : LENGTH: 1815
      :
      : TYPE: DNA
      :
      : ORGANISM: Artificial Sequence
      : FEATURE:
      :
      : OTHER INFORMATION: Synthetic B.t. toxin gene
      :
      : US-09-178-252-8

```

	OY	3	TGAACTCCTGTATCGCATTCATATTGGGGAGTCCGAATGCCACGCACAT	53			
D8	430	TGATGGCGCGTGATCAAGACGTCATCGTTGGTAGAAGGATTAACAACAT	380				
	Matches	35;	Conservative	0; Mismatches	16; Indels	0; Gaps	0;
	Query Match	Score	25.4;	DB	4;	length	1815;
	Best Local Similarity	Pred.	No. 2.1;				
			68.6%;				

RESULT 8
US-09-038-832-3/c
Sequence 3, Application US/09038832
Patent No. 6146845
GENERAL INFORMATION:
APPLICANT: KIKLY, KRISTINE
APPLICANT: ERICKSON-MILLER, CONNIE
TITLE OF INVENTION: Sialodhesin Family Member-2
TITLE OF INVENTION: (SAF-2)
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RAINIER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/038,832
FILING DATE: 11-MAR-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/041,886
FILING DATE: 02-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-50018
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700

Db 4427 TGACATTTTTCATTCCTTTTCTTGTGGAATCATTAATGAGACACATTATCAAA 4486
 QY 63 ATCCCACTTGACCAACTGAAGAAA 88
 Db 4487 ATACCAATGTATCAAAATGAGTAAA 4512

RESULT 2

US-08-370-975B-1
 ; Sequence 1, Application US/08370975B
 ; Patent No. 5622851

GENERAL INFORMATION:

APPLICANT: Maley, Frank
 APPLICANT: Maley, Gladys F.
 APPLICANT: Weiner, Karen X.B.
 TITLE OF INVENTION: Human Deoxycytidylate Deaminase Gene
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Nixon, Hargrave, Devans & Doyle
 STREET: Clinton Square, P.O. Box 1051
 CITY: Rochester
 STATE: New York
 COUNTRY: USA
 ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA: US/08/370,975B
 APPLICATION NUMBER: US/08/370,975B
 FILING DATE: 10-JAN-1995
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Timlan, Susan J.
 REGISTRATION NUMBER: 34,103
 REFERENCE/DOCKET NUMBER: 20894/80
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (716)263-1636
 TELEFAX: (716)263-1600

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 26764 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT: 4q35
 US-08-370-975B-1

Query Match 29.8%; Score 26.8; DB 1; Length 26764;
 Best Local Similarity 57.0%; Pred. No. 1.7;
 Matches 49; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 3 TGACCTCTGTGATCGAATCAATCGGATGCGGATGGGACGACATAAATCCAA 62
 Db 6390 TGACATTTTTCATTCCTTTTCTTGTGGAATCATTAATGAGACACATTATCAAA 4449
 QY 63 ATCCCACTTGACCAACTGAAGAAA 88
 Db 6450 ATACCAATGTATCAAAATGAGTAAA 4475

RESULT 3

US-09-484-970B-93
 ; Sequence 93, Application US/09484970B
 ; Patent No. 6426186

GENERAL INFORMATION:

APPLICANT: Jones, Karen A.
 APPLICANT: Volkmuth, Wayne
 APPLICANT: Walker, Michael G.
 TITLE OF INVENTION: BONE REMODELING GENES

FILE REFERENCE: PB-0014 US
 CURRENT APPLICATION NUMBER: US/09/484,970B
 CURRENT FILING DATE: 2000-01-18
 NUMBER OF SEQ ID NOS: 172
 SOFTWARE: PERL Program
 SEQ ID NO 93
 LENGTH: 3446
 TYPE: DNA

ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc.feature
 OTHER INFORMATION: Incyte ID No. 6426186 252234.3CB1
 NAME/KEY: unsure
 LOCATION: 1432, 1435
 OTHER INFORMATION: a, t, c, g, or other
 US-09-484-970B-93

Query Match 28.9%; Score 26; DB 4; Length 3446;
 Best Local Similarity 76.2%; Pred. No. 1.6;
 Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 48 GCACATTAATCCCAATCCCACTTGAACCACTGAAGAAAAG 89
 Db 775 GACCTAAATCTCTGTCACCACTGAGCCAGCAAGAGAGAG 816

RESULT 4

US-09-283-763-1
 ; Sequence 1, Application US/09283763
 ; Patent No. 6110685

GENERAL INFORMATION:

APPLICANT: Zalacain, Magdalena
 APPLICANT: Burnham, Martin K. R.
 APPLICANT: Biswas, Sanjoy
 APPLICANT: Brown, James R.
 APPLICANT: Ingraham, Karen A.
 APPLICANT: Chaliker, Allison F.
 APPLICANT: So, Chi Young
 APPLICANT: Holmes, David J.
 APPLICANT: Traini, Christopher M.
 APPLICANT: Warren, Richard L.
 APPLICANT: Mathie, Thomas B.
 TITLE OF INVENTION: Infb
 FILE REFERENCE: GM10175
 CURRENT APPLICATION NUMBER: US/09/283,763
 CURRENT FILING DATE: 1999-04-01
 EARLIER APPLICATION NUMBER: 60/105,985
 EARLIER FILING DATE: 1999-10-28
 NUMBER OF SEQ ID NOS: 3
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 1
 LENGTH: 2793
 TYPE: DNA
 ORGANISM: Streptococcus pneumoniae
 US-09-283-763-1

Query Match 28.7%; Score 25.8; DB 3; Length 2793;
 Best Local Similarity 56.5%; Pred. No. 1.7;
 Matches 48; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 2 GTGACCTCTGTGATCGAATCATTCATTCGCGGATGGGAGGAGCGACATAAATCCAA 61
 Db 2615 GTGTATCCGTGATGAGGTGTCTATCTATGATGCGGACACTGCAAGCTTGAACACTACA 2674

QY 62 AATCCACTTGACCAACTGAAGAA 86
 Db 2675 AAGATGACGTGAAGAAAGTGAACAA 2699

RESULT 5

US-09-574-912-1
 ; Sequence 1, Application US/09574912
 ; Patent No. 6399343

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:23:43 ; Search time 4.30784 Seconds

(without alignments)
6407.139 Million cell updates/sec

Title: US-09-513-151-3_COPY_1121_1210

Perfect score: 90
Sequence: 1 TGTGACCTCTGTCGATCGAAT.....TGACCACTGAGAGAAACA 90

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26.8	29.8	20303	1 US-08-370-975B-6	Sequence 6, App1
2	26.8	29.8	26764	1 US-08-370-975B-1	Sequence 1, App1
3	26	28.9	3446	4 US-09-484-970B-93	Sequence 93, App1
4	25.8	28.7	2793	3 US-09-283-763-1	Sequence 1, App1
5	25.8	28.7	2793	4 US-09-574-912-1	Sequence 1, App1
6	25.4	28.2	1815	4 US-09-178-252-7	Sequence 7, App1
7	25.4	28.2	1815	4 US-09-178-252-8	Sequence 8, App1
8	25.2	28.0	1597	3 US-09-338-832-3	Sequence 3, App1
9	25.2	28.0	2900	3 US-09-338-832-1	Sequence 1, App1
10	24.2	26.9	2787	4 US-09-277-565-27	Sequence 27, App1
11	24.2	26.9	11378	4 US-08-961-827-210	Sequence 210, App1
12	24	26.7	4629	2 US-08-484-891-7	Sequence 1, App1
13	24	26.7	4670	3 US-08-717-294-41	Sequence 41, App1
14	24	26.7	4999	4 US-09-470-618-14	Sequence 14, App1
15	24	26.7	4999	4 US-09-364-862-14	Sequence 14, App1
16	24	26.7	5035	2 US-08-882-083-3	Sequence 1, App1
17	24	26.7	5035	2 US-08-558-107-1	Sequence 1, App1
18	24	26.7	5035	3 US-09-243-539-1	Sequence 1, App1
19	24	26.7	6999	1 US-08-276-594A-1	Sequence 1, App1
20	24	26.7	7056	1 US-08-121-202-1	Sequence 1, App1
21	24	26.7	8241	6 5171844-1	Patent No. 5171844
22	24	26.7	8967	1 US-08-366-851A-1	Sequence 1, App1
23	24	26.7	9009	1 US-07-864-004B-3	Sequence 3, App1
24	24	26.7	9009	1 US-08-251-937A-3	Sequence 3, App1
25	24	26.7	9009	1 US-08-212-133A-1	Sequence 1, App1
26	24	26.7	9009	1 US-08-474-503-1	Sequence 1, App1
27	24	26.7	9009	2 US-08-670-707A-1	Sequence 1, App1

C 28	24	26.7	9009	4 US-09-037-601-1	Sequence 1, App1
C 29	24	26.7	9009	4 US-09-315-179-1	Sequence 1, App1
C 30	24	26.7	9009	4 US-09-523-656-1	Sequence 1, App1
C 31	24	26.7	9009	5 PCT-US93-03275-3	Sequence 3, App1
C 32	24	26.7	9009	5 PCT-US94-13200-1	Sequence 1, App1
C 33	24	26.7	9354	1 US-08-683-839B-2	Sequence 2, App1
C 34	24	26.7	11933	4 US-09-470-618-13	Sequence 13, App1
C 35	24	26.7	11933	4 US-09-364-862-13	Sequence 13, App1
C 36	24	26.7	11933	4 US-09-178-252-3	Sequence 3, App1
C 37	23.8	26.4	1815	4 US-09-178-252-1	Sequence 1, App1
C 38	23.8	26.4	3444	4 US-09-853-913-1	Sequence 1, App1
C 39	23.6	26.2	4326	2 US-08-852-807-12	Sequence 12, App1
C 40	23.6	26.2	13674	2 US-08-852-807-1	Sequence 1, App1
C 41	23.4	26.0	169998	4 US-09-676-610B-24	Sequence 24, App1
C 42	23.2	25.8	1033	4 US-09-171-209-73	Sequence 73, App1
C 43	23.2	25.8	1662	3 US-08-811-177A-1	Sequence 1, App1
C 44	23.2	25.8	55827	4 US-09-813-133A-3	Sequence 3, App1
C 45	23.2	25.8	168575	4 US-09-426-290-1	Sequence 1, App1

ALIGNMENTS

```
RESULT 1
US-08-370-975B-6
Sequence 6, Application US/08370975B
Patent No. 5622851
GENERAL INFORMATION:
APPLICANT: Maley, Frank
APPLICANT: Weiner, Karen X.B.
TITLE OF INVENTION: Human Deoxycytidylate Deaminase Gene
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,975B
FILING DATE: 10-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Timlan, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 20894/80
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716)263-1636
TELEFAX: (716)263-1600
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 20303 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 4q35
US-08-370-975B-6
Query Match 29.8%; Score 26.8; DB 1; Length 20303;
Best Local Similarity 57.0%; Pred. No. 1.5;
Matches 49; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
3 TGTGACCTCTGTCGATCTGTCGATCGAATGGGACCGCACATTAATCA 62
|||||
```

```

; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1128
; LENGTH: 653
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-1128

```

```

Query Match      27.6%; Score 24.8; DB 10; Length 653;
Best Local Similarity 63.3%; Pred. No. 14;
Matches 38; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

```

QY      4 GACCTCTGTGATCATCATTTGGGATCGGATGGGACGCGACATTAATCCANA 63
DB      268 GACCTTAGAGACCTTATATCATGTGTGAGAGCGAAGGGGAGCAACAGAGAGANA 327

```

RESULT 15

```

US-09-925-302-212
; Sequence 212, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 212
; LENGTH: 1529
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-212

```

```

Query Match      27.6%; Score 24.8; DB 10; Length 1529;
Best Local Similarity 56.0%; Pred. No. 20;
Matches 47; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

```

QY      7 CTCTGTGATCGAATCATCATTTGGGATCGGATGGGACGCGACATTAATCCAAATCC 66
DB      403 CTCTGCTCTCCAAAGCCCTTCTCATGTGTGGGAGGTGGAAGGACATTAAGGAATCCC 462

```

```

QY      67 CACTTGAAACCACTGAAGAAAGA 90
DB      463 CATTTGAGAGTACAGAAAGAAAGA 486

```

Search completed: April 21, 2003, 18:47:45
Job time : 13.3895 secs

RESULT 12
US-09-972-714-9/c

RESULT 14
US-09-764-860-1128
Sequence 1128, Application US/09764860.
Patent No. US20020094953A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.,
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

QY	13	AA	14
		11	
Db	22	AA	21


```
Sequence 9565, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haseldeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9565
LENGTH: 2703
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2703)
US-09-815-242-9565

Query Match          30.4%; Score 27.4; DB 10; Length 2703;
Best Local Similarity 57.6%; Pred. No. 2.6;
Matches 49; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 2 GTGACCTGTGATGGAATCATTCATTGGGATCGCGATGCGCAGCAGCATTAATCCA 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2525 GTGTATCGGTGATGCGTGTATCTATGATGCGGAGTCCAGCTTGAACACTACA 2584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 62 AATCCACTTGAAACCACTGAAGAA 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2585 AAGACGACGTGAAGAAGTGAACAAA 2609

RESULT 3
US-09-880-192-41
Sequence 41, Application US/09880192
Patent No. US20020077470A1
GENERAL INFORMATION:
APPLICANT: Walker, Michael G.
APPLICANT: Klingler, Tod M.
APPLICANT: Azimzal, Yalda
TITLE OF INVENTION: POLYNUCLEOTIDES ASSOCIATED WITH CARDIAC MUSCLE FUNCTION
FILE REFERENCE: PB-0009-1 CIP
CURRENT APPLICATION NUMBER: US/09/880,192
CURRENT FILING DATE: 2001-06-12
NUMBER OF SEQ ID NOS: 62
SOFTWARE: PERL Program
SEQ ID NO 41
LENGTH: 3441
TYPE: DNA
ORGANISM: Homo sapiens
```

```
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020077470A1 4151935CB1
US-09-880-192-41

Query Match          28.9%; Score 26; DB 10; Length 3441;
Best Local Similarity 76.2%; Pred. No. 9.4;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 48 GCACATAAATCCAAATCCCACTTGAAACCACTGAAGAAG 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 775 GAACCTAAATCTGTCCTCCACTGAGCCAGTGAAGAGAG 816

RESULT 4
US-10-114-893-126
Sequence 126, Application US/10114893
Publication No. US20020193567A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Bowman, Michael R.
APPLICANT: Spaulding, Vikki
APPLICANT: Carlin-Duckett, McKeough
APPLICANT: Kelleher, Kerry S.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: GI 6000-10A
CURRENT APPLICATION NUMBER: US/10/114,893
CURRENT FILING DATE: 2002-04-02
EARLIER APPLICATION NUMBER: 09/413,232
EARLIER FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 321
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 126
LENGTH: 3444
TYPE: DNA
ORGANISM: Homo sapiens
US-10-114-893-126

Query Match          28.9%; Score 26; DB 9; Length 3444;
Best Local Similarity 76.2%; Pred. No. 9.4;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 48 GCACATAAATCCAAATCCCACTTGAAACCACTGAAGAAG 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 809 GAACCTAAATCTGTCCTCCACTGAGCCAGTGAAGAGAG 850

RESULT 5
US-10-091-504-2256
Sequence 2256, Application US/10091504
Publication No. US20030059908A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC007C1
CURRENT APPLICATION NUMBER: US/10/091,504
CURRENT FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 2442
Prior Application removed - See File Wrapper or Palm
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2256
LENGTH: 4346
TYPE: DNA
ORGANISM: Homo sapiens
US-10-091-504-2256
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 15:55:39 ; Search time 10.3895 Seconds
(without alignments)
8702.120 Million cell updates/sec

Title: US-09-513-151-3_COPY_1121_1210

Perfect score: 90

Sequence: 1 TGTGACCTCTGTGATCGAAT.....TGACCACTGAGAGAAAAGA 90

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PC1_NEM_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PC1_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27.4	30.4	2000	US-09-938-842A-3043	Sequence 3043, Ap
2	27.4	30.4	2703	US-09-815-242-9565	Sequence 9565, Ap
3	26	28.9	3441	US-09-880-182-41	Sequence 41, Appl
4	26	28.9	3444	US-10-114-893-126	Sequence 126, Ap
5	26	28.9	4346	US-10-091-504-2256	Sequence 2256, Ap
6	26	28.9	4346	US-09-764-869-2256	Sequence 2256, Ap
7	25.4	28.2	1815	US-09-826-660-7	Sequence 7, Appl
8	25.4	28.2	1815	US-08-826-660-8	Sequence 8, Appl
9	25.2	28.0	182	US-09-969-373-1375	Sequence 1375, Ap
10	25.2	28.0	502	US-09-972-714-2	Sequence 2, Appl
11	25.2	28.0	1554	US-09-910-600-5	Sequence 5, Appl
12	25.2	28.0	1592	US-09-972-714-9	Sequence 9, Appl
13	25.2	28.0	6101	US-09-972-714-1	Sequence 1, Appl
14	24.8	27.6	653	US-09-764-860-1128	Sequence 1128, Ap
15	24.8	27.6	1529	US-09-925-302-212	Sequence 212, Ap
16	24.6	27.3	5470	US-09-764-864-1593	Sequence 1593, Ap
17	24.4	27.1	483	US-09-943-718-1	Sequence 1, Appl
18	24.4	27.1	912	US-09-792-793A-70	Sequence 70, Appl
19	24.4	27.1	1042	US-10-066-500-77	Sequence 77, Appl

C 20	24.4	27.1	1042	9	US-10-002-796-77	Sequence 77, Appl
C 21	24.4	27.1	1042	9	US-10-066-273-77	Sequence 77, Appl
C 22	24.4	27.1	1042	9	US-10-066-494-77	Sequence 77, Appl
C 23	24.4	27.1	1042	9	US-10-066-269-77	Sequence 77, Appl
C 24	24.4	27.1	1042	9	US-10-066-193-77	Sequence 77, Appl
C 25	24.4	27.1	1042	9	US-10-066-211-77	Sequence 77, Appl
C 26	24.2	26.9	1248	9	US-09-938-842A-266	Sequence 266, Ap
C 27	24.2	26.7	4629	10	US-09-150-811-7	GENERAL INFO
C 28	24.2	26.7	4999	9	US-10-007-968-14	Sequence 14, Appl
C 29	24.2	26.7	4999	10	US-09-740-211-14	Sequence 14, Appl
C 30	24.2	26.7	7944	12	US-10-095-718-1	Sequence 1, Appl
C 31	24.2	26.7	9009	9	US-09-957-641-1	Sequence 1, Appl
C 32	24.2	26.7	9009	9	US-10-187-319-1	Sequence 1, Appl
C 33	24.2	26.7	9029	9	US-10-132-828-1	Sequence 1, Appl
C 34	24.2	26.7	11933	9	US-10-007-968-13	Sequence 13, Appl
C 35	24.2	26.7	11933	10	US-09-740-211-13	Sequence 13, Appl
C 36	23.8	26.4	396	10	US-09-960-352-11737	Sequence 11737, A
C 37	23.8	26.4	462	9	US-09-925-299-335	Sequence 335, Ap
C 38	23.8	26.4	462	10	US-09-925-289-335	Sequence 335, Ap
C 39	23.8	26.4	818	9	US-10-202-193-116	Sequence 116, Appl
C 40	23.8	26.4	1037	9	US-09-974-879-53	Sequence 53, Appl
C 41	23.8	26.4	1815	10	US-09-826-660-3	Sequence 3, Appl
C 42	23.8	26.4	3444	10	US-09-826-660-1	Sequence 1, Appl
C 43	23.6	26.2	424	10	US-09-983-965-4039	Sequence 4039, Ap
C 44	23.6	26.2	858	9	US-09-738-628-3057	Sequence 3057, Ap
C 45	23.6	26.2	2061	9	US-09-938-842A-1375	Sequence 1375, Ap

ALIGNMENTS

RESULT 1
US-09-938-842A-3043
Sequence 3043, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kleps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938, 842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227, 866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264, 647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300, 111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 3043
LENGTH: 2000
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-3043
Query Match
Best Local Similarity 30.4%; Score 27.4; DB 9; Length 2000;
Matches 49; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 2 GTGACCTCTGTGATCGAATCATTCATTTGGGATCGCAATGGCGAGCGACATATAATCA 61
DB 337 GAGCCATCGAGTCTGATCATTCATTTAGCCGCGACATGCGACCGACCATTCCTTAACA 396
QY 62 AATCCACTGGAACGACGAGAA 86
DB 397 CATCTCTTCACGCGAGATGACAA 421
RESULT 2
US-09-815-242-9565

Query Match 100.0%; Score 90; DB 17; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.5e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGACGCCACATAAATCC 60
|||||
DB 923 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGACGCCACATAAATCC 864
OY 61 AATCCCACTTGAAACCACTGAAGAAAGA 90
|||||
DB 863 AATCCCACTTGAAACCACTGAAGAAAGA 834

RESULT 14

US-09-898-888-8140/c
; Sequence 8140, Application US/09898888
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-748CON1
; CURRENT APPLICATION NUMBER: US/09/898,888
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 09/340,623
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 45207
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 8140
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-898-888-8140

Query Match 100.0%; Score 90; DB 33; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.5e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGACGCCACATAAATCC 60
|||||
DB 923 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGACGCCACATAAATCC 864
OY 61 AATCCCACTTGAAACCACTGAAGAAAGA 90
|||||
DB 863 AATCCCACTTGAAACCACTGAAGAAAGA 834

RESULT 15

US-09-898-888A-8140/c
; Sequence 8140, Application US/09898888A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-748CON1
; CURRENT APPLICATION NUMBER: US/09/898,888A
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/340,623
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/205,070
; PRIOR FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 45207
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 8140
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-898-888A-8140

Query Match 100.0%; Score 90; DB 33; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.5e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGACGCCACATAAATCC 60
|||||
DB 923 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGACGCCACATAAATCC 864
OY 61 AATCCCACTTGAAACCACTGAAGAAAGA 90
|||||
DB 863 AATCCCACTTGAAACCACTGAAGAAAGA 834

Search completed: April 21, 2003, 17:14:11
Job time : 267.819 secs

;; CURRENT FILING DATE: 1999-07-22
;; EARLIER APPLICATION NUMBER: US 09/205,155
;; EARLIER FILING DATE: 1998-12-03
;; NUMBER OF SEQ ID NOS: 13203
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 3465
;; LENGTH: 1039
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-359-922-3465

Query Match 100.0%; Score 90; DB 17; Length 1039;
Best Local Similarity 100.0%; Pred. No. 1.5e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGACACATAAAATCC 60
|||||
DB 921 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGACACATAAAATCC 862
|||||
OY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
|||||
DB 861 AATCCCACTTGAACCACTGAAGAAAAGA 832

RESULT 10
US-09-359-922-3465/c
;; Sequence 3465, Application US/09359922A
;; GENERAL INFORMATION:
;; APPLICANT: Leshkowitz, Dena
;; APPLICANT: Liu, Jin
;; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
;; FILE REFERENCE: 20411-752CON1
;; CURRENT APPLICATION NUMBER: US/09/359,922A
;; CURRENT FILING DATE: 1999-07-22
;; EARLIER APPLICATION NUMBER: US 09/205,155
;; EARLIER FILING DATE: 1998-12-03
;; EARLIER APPLICATION NUMBER: US 09/034,341
;; EARLIER FILING DATE: 1998-02-13
;; NUMBER OF SEQ ID NOS: 13203
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 3465
;; LENGTH: 1039
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-359-922-3465

Query Match 100.0%; Score 90; DB 17; Length 1039;
Best Local Similarity 100.0%; Pred. No. 1.5e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGACACATAAAATCC 60
|||||
DB 921 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGACACATAAAATCC 862
|||||
OY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
|||||
DB 861 AATCCCACTTGAACCACTGAAGAAAAGA 832

RESULT 11
US-09-919-002-3465/c
;; Sequence 3465, Application US/09919002
;; GENERAL INFORMATION:
;; APPLICANT: Leshkowitz, Dena
;; APPLICANT: Liu, Jin
;; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
;; FILE REFERENCE: 20411-752CON1
;; CURRENT APPLICATION NUMBER: US/09/919,002
;; CURRENT FILING DATE: 2001-07-30
;; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 09/359,922
;; PRIOR FILING DATE: FILING DATE: 1999-07-22

;; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US 09/034,341
;; PRIOR FILING DATE: FILING DATE: 1998-02-13
;; NUMBER OF SEQ ID NOS: 13203
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 3465
;; LENGTH: 1039
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-919-002-3465

Query Match 100.0%; Score 90; DB 34; Length 1039;
Best Local Similarity 100.0%; Pred. No. 1.5e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGACACATAAAATCC 60
|||||
DB 921 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGACACATAAAATCC 862
|||||
OY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
|||||
DB 861 AATCCCACTTGAACCACTGAAGAAAAGA 832

RESULT 12
US-09-205-070-8140/c
;; Sequence 8140, Application US/09205070
;; GENERAL INFORMATION:
;; APPLICANT: Hysq, Inc.
;; APPLICANT: Hysq, Inc.
;; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
;; FILE REFERENCE: 20411-748
;; CURRENT APPLICATION NUMBER: US/09/205,070
;; CURRENT FILING DATE: 1998-12-03
;; NUMBER OF SEQ ID NOS: 45207
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 8140
;; LENGTH: 1041
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-205-070-8140

Query Match 100.0%; Score 90; DB 16; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.5e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGACACATAAAATCC 60
|||||
DB 923 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGACACATAAAATCC 864
|||||
OY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
|||||
DB 863 AATCCCACTTGAACCACTGAAGAAAAGA 834

RESULT 13
US-09-340-623-8140/c
;; Sequence 8140, Application US/09340623
;; GENERAL INFORMATION:
;; APPLICANT: Hysq, Inc.
;; APPLICANT: Hysq, Inc.
;; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
;; FILE REFERENCE: 20411-748CON1
;; CURRENT APPLICATION NUMBER: US/09/340,623
;; CURRENT FILING DATE: 1999-06-28
;; EARLIER APPLICATION NUMBER: US 09/205,070
;; EARLIER FILING DATE: 1998-12-03
;; NUMBER OF SEQ ID NOS: 45207
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 8140
;; LENGTH: 1041
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-340-623-8140

```
OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGATGGCGACGACATAAATCC 60
    |||
DB 278 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGATGGCGACGACATAAATCC 337
OY 61 AAATCCCACTTGACCACTGAGAAAAGA 90
    |||
DB 338 AAATCCCACTTGACCACTGAGAAAAGA 367
```

RESULT 7

```
US-09-515-128-21383
; Sequence 21383, Application US/09515128
; GENERAL INFORMATION:
```

```
APPLICANT: Atterburn, Matthew
APPLICANT: Asghari, Vida
APPLICANT: Damavandi, Simin
APPLICANT: Dickson, Mark
APPLICANT: Drake, Jim
APPLICANT: Drmanac, Radoje
APPLICANT: Engleman, Carrie
APPLICANT: Faulkner, Brandy
APPLICANT: Garcia, Veronica
APPLICANT: Giedt, Gretchen
APPLICANT: Hunter, Kelly
APPLICANT: Jessen, Aaron
APPLICANT: Jones, Lee
APPLICANT: Kita, David
APPLICANT: Labat, Ivan
APPLICANT: Laroza, Mimi
APPLICANT: Lomelli, Michelle
APPLICANT: Nguyen, Phuong
APPLICANT: Nogra, Margie
APPLICANT: Palencia, Servando
APPLICANT: Ralsi, Fariba
APPLICANT: Smith, Benjamin
APPLICANT: Tkach, Joe
APPLICANT: Tran, Lien
APPLICANT: Verna, Ron
APPLICANT: Yang, Fel
APPLICANT: Yim, Kenneth
TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
FILE REFERENCE: 728CIP
CURRENT APPLICATION NUMBER: US/09/515,128
CURRENT FILING DATE: 2000-02-28
EARLIER APPLICATION NUMBER: 09/034,341
EARLIER FILING DATE: 1998-02-13
NUMBER OF SEQ ID NOS: 24489
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 21383
LENGTH: 442
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(442)
OTHER INFORMATION: n = A,T,C or G
```

```
US-09-515-128-21383
```

```
Query Match 100.0%; Score 90; DB 19; Length 442;
Best Local Similarity 100.0%; Pred. No. 1.2e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGACGACATAAATCC 60
    |||
DB 149 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGACGACATAAATCC 208
OY 61 AAATCCCACTTGACCACTGAGAAAAGA 90
    |||
DB 209 AAATCCCACTTGACCACTGAGAAAAGA 238
```

RESULT 8

```
US-09-721-544-21383
```

```
; Sequence 21383, Application US/09721544
```

```
; GENERAL INFORMATION:
```

```
APPLICANT: Atterburn, Matthew
APPLICANT: Asghari, Vida
APPLICANT: Damavandi, Simin
APPLICANT: Dickson, Mark
APPLICANT: Drake, Jim
APPLICANT: Drmanac, Radoje
APPLICANT: Engleman, Carrie
APPLICANT: Faulkner, Brandy
APPLICANT: Garcia, Veronica
APPLICANT: Giedt, Gretchen
APPLICANT: Hunter, Kelly
APPLICANT: Jessen, Aaron
APPLICANT: Jones, Lee
APPLICANT: Kita, David
APPLICANT: Labat, Ivan
APPLICANT: Laroza, Mimi
APPLICANT: Lomelli, Michelle
APPLICANT: Nguyen, Phuong
APPLICANT: Nogra, Margie
APPLICANT: Palencia, Servando
APPLICANT: Ralsi, Fariba
APPLICANT: Smith, Benjamin
APPLICANT: Tkach, Joe
APPLICANT: Tran, Lien
APPLICANT: Verna, Ron
APPLICANT: Yang, Fel
APPLICANT: Yim, Kenneth
TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
FILE REFERENCE: 728CIP
CURRENT APPLICATION NUMBER: US/09/721,544
CURRENT FILING DATE: 2000-11-21
PRIOR APPLICATION NUMBER: 09/515,128
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: 09/034,341
PRIOR FILING DATE: 1998-02-13
NUMBER OF SEQ ID NOS: 24489
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 21383
LENGTH: 442
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(442)
OTHER INFORMATION: n = A,T,C or G
```

```
US-09-721-544-21383
```

```
Query Match 100.0%; Score 90; DB 29; Length 442;
Best Local Similarity 100.0%; Pred. No. 1.2e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGACGACATAAATCC 60
    |||
DB 149 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGACGACATAAATCC 208
OY 61 AAATCCCACTTGACCACTGAGAAAAGA 90
    |||
DB 209 AAATCCCACTTGACCACTGAGAAAAGA 238
```

RESULT 9

```
US-09-359-922-3465/C
; Sequence 3465, Application US/09359922
```

```
; GENERAL INFORMATION:
```

```
APPLICANT: HySeq, Inc.
TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS cDNA
FILE REFERENCE: 20411-752CON1
CURRENT APPLICATION NUMBER: US/09/359,922
```

```
APPLICANT: Naughton, Rebecca E.
APPLICANT: Klingler, Tod M.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
TITLE OF INVENTION: HUMAN FETAL LUNG
NUMBER OF SEQUENCES: 2102
CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/044,082
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PD-0357P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 268:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 2888829H1
US-60-044-082-268

Query Match
Best Local Similarity 100.0%; Score 90; DB 48; Length 263;
Pred. No. 1e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTACCTCTGTGATCATCATTTGGGGATGCGGATGGGAGCGGCACATAAAATCC 60
DB 34 TGTACCTCTGTGATCATCATTTGGGGATGCGGATGGGAGCGGCACATAAAATCC 93
QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
DB 94 AATCCCACTTGAACCACTGAAGAAAAGA 123

RESULT 4
US-09-532-315-9519
Sequence 9519, Application US/09532315
GENERAL INFORMATION:
APPLICANT: Sellhammer, Jeffrey J.
APPLICANT: Deleageane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullahy, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES
FILE REFERENCE: PD-1002 CIP
CURRENT APPLICATION NUMBER: US/09/532,315
CURRENT FILING DATE: 2000-03-24
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 42212
SOFTWARE: PERL Program
SEQ ID NO 9519
LENGTH: 289
TYPE: DNA
ORGANISM: Homo sapiens
```

```
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: hu00815059
US-09-532-315-9519

Query Match
Best Local Similarity 100.0%; Score 90; DB 20; Length 289;
Pred. No. 1e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTACCTCTGTGATCATCATTTGGGGATGCGGATGGGAGCGGCACATAAAATCC 60
DB 84 TGTACCTCTGTGATCATCATTTGGGGATGCGGATGGGAGCGGCACATAAAATCC 143
QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
DB 144 AATCCCACTTGAACCACTGAAGAAAAGA 173

RESULT 5
US-09-489-036-7332
Sequence 7332, Application US/09489036
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
FILE REFERENCE: 783
CURRENT APPLICATION NUMBER: US/09/489,036
CURRENT FILING DATE: 2000-01-19
NUMBER OF SEQ ID NOS: 35324
SOFTWARE: HY-patent.pl Version 3.1
SEQ ID NO 7332
LENGTH: 372
TYPE: DNA
ORGANISM: Homo sapiens
US-09-489-036-7332

Query Match
Best Local Similarity 100.0%; Score 90; DB 18; Length 372;
Pred. No. 1e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTACCTCTGTGATCATCATTTGGGGATGCGGATGGGAGCGGCACATAAAATCC 60
DB 278 TGTACCTCTGTGATCATCATTTGGGGATGCGGATGGGAGCGGCACATAAAATCC 337
QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
DB 338 AATCCCACTTGAACCACTGAAGAAAAGA 367

RESULT 6
US-09-943-143-7332
Sequence 7332, Application US/09943143
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
FILE REFERENCE: 783
CURRENT APPLICATION NUMBER: US/09/943,143
CURRENT FILING DATE: 2001-08-30
Prior application data removed - refer to PALM or file wrapper
Prior FILING DATE: 2000-01-19
NUMBER OF SEQ ID NOS: 35324
SOFTWARE: HY-patent.pl Version 3.1
SEQ ID NO 7332
LENGTH: 372
TYPE: DNA
ORGANISM: Homo sapiens
US-09-943-143-7332

Query Match
Best Local Similarity 100.0%; Score 90; DB 35; Length 372;
Pred. No. 1e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

22 90 100.0 1858 18 US-09-471-275-2814
23 90 100.0 2041 19 US-09-513-151-3
24 90 100.0 2101 61 US-60-172-360-26034
25 90 100.0 2108 40 US-10-133-013-165
26 90 100.0 2109 76 US-60-324-185-33690
27 90 100.0 2129 1 PCR-US02-07826-150
28 90 100.0 2129 1 PCR-US02-07826-150
29 90 100.0 2129 39 US-10-097-340-152
30 90 100.0 2129 39 US-10-097-340-152
31 88.4 98.2 1069 25 US-09-652-124-7628
32 75.4 83.8 287 12 US-08-810-326-938
33 75.4 83.8 287 20 US-09-532-315-9504
34 75.4 83.8 287 45 US-60-012-699-938
35 74 82.2 2103 80 US-60-360-207-2119
36 72.4 80.4 301 15 US-09-151-199-3234
37 72.4 80.4 301 18 US-09-482-965-1320
38 72.4 80.4 301 23 US-09-539-806-4638
39 56.4 62.7 523 20 US-09-617-081-981
40 52.8 58.7 434 11 US-08-798-074-1079
41 52.8 58.7 434 11 US-08-798-074B-1079
42 49.4 54.9 51 76 US-60-324-185-33691
43 47.2 52.4 221 35 US-09-880-093-760
44 47.2 52.4 248 13 US-09-107-425-1617
45 47.2 52.4 248 21 US-09-540-764-25285

Sequence 2814, Ap
Sequence 3, Appli
Sequence 26034, A
Sequence 165, App
Sequence 33690, A
Sequence 150, App
Sequence 150, App
Sequence 152, App
Sequence 152, App
Sequence 7628, Ap
Sequence 938, App
Sequence 9504, App
Sequence 938, App
Sequence 2119, Ap
Sequence 3234, Ap
Sequence 1320, Ap
Sequence 4638, A
Sequence 981, App
Sequence 1079, Ap
Sequence 1079, Ap
Sequence 33691, A
Sequence 760, App
Sequence 1617, Ap
Sequence 25285, A

ALIGNMENTS

RESULT 1

US-09-056-942-268

Sequence 268, Application US/09056942

GENERAL INFORMATION:

APPLICANT: Gooding, Douglas H.

APPLICANT: Stuve, Laura L.

APPLICANT: Stuart, Susan G.

APPLICANT: Ito, Laura Y.

APPLICANT: Akerblom, Ingrid E.

APPLICANT: Delegeane, Angelo M.

APPLICANT: Naughton, Rebecca E.

APPLICANT: Klingler, Tod M.

TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM

NUMBER OF SEQUENCES: 2102

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/056,942

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/044,082

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: GEROME, MICHAEL C.

REGISTRATION NUMBER: 39,132

REFERENCE/DOCKET NUMBER: PD-0357P

TELEPHONE: (415) 855-0555

TELEFAX: (415) 845-4166

INFORMATION FOR SEQ ID NO: 268:

SEQUENCE CHARACTERISTICS:

LENGTH: 263 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 2888829H1

US-09-056-942-268

Query Match 100.0%; Score 90; DB 14; Length 263;

Best Local Similarity 100.0%; Pred. No. 1e-21; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:26:48 ; Search time 265.819 Seconds

(without alignments)
8512.654 Million cell updates/sec

Title: US-09-513-151-3_COPY_1121_1210
Perfect score: 90
Sequence: 1 TGTGACCTGTGTGATCGAAT.....TGAACCACTGAGAGAAAAGA 90

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 08
Listing first 45 summaries

Database :

Pending Patents_NA_Main:*

```

1: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US0958_COMB.seq.*

```

Result		Query		DB		ID		Description	
No.	Score	Match	Length						
1	90	100.0	263	14	US-09-056-942-268			Sequence 268, App	
2	90	100.0	263	20	US-09-532-315-9510			Sequence 9510, App	
3	90	100.0	263	48	US-60-044-082-268			Sequence 268, App	
4	90	100.0	289	20	US-09-532-315-9519			Sequence 9519, App	
5	90	100.0	372	18	US-09-489-036-7332			Sequence 7332, App	
6	90	100.0	372	35	US-09-943-143-7332			Sequence 7332, App	
7	90	100.0	442	19	US-09-515-188-21383			Sequence 21383, A	
8	90	100.0	442	29	US-09-721-544-21383			Sequence 21383, A	
9	90	100.0	1039	17	US-09-359-922-3465			Sequence 3465, App	
10	90	100.0	1039	17	US-09-359-922-3465			Sequence 3465, App	
11	90	100.0	1039	34	US-09-919-002-8146			Sequence 8146, App	
12	90	100.0	1041	16	US-09-205-070-8140			Sequence 8140, App	
13	90	100.0	1041	17	US-09-340-623-8140			Sequence 8140, App	
14	90	100.0	1041	33	US-09-898-888-8140			Sequence 8140, App	
15	90	100.0	1041	33	US-09-898-888-8140			Sequence 8140, App	
16	90	100.0	1183	42	US-09-757-028-874			Sequence 874, App	
17	90	100.0	1183	29	US-10-232-911-874			Sequence 874, App	
18	90	100.0	1223	28	US-09-705-256A-6006			Sequence 6006, App	
19	90	100.0	1223	60	US-60-164-285-6006			Sequence 6006, App	
20	90	100.0	1749	1	PCT-US02-18947-1692			Sequence 1692, App	
21	90	100.0	1749	41	US-10-172-118-1692			Sequence 1692, App	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Db 1211 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCGATGGCAGCGACATATAAATCC 12710
OY 61 AATCCCACTTGAACCACTGAGAGAAAAGA 90
Db 1271 AATCCCACTTGAACCACTGAGAGAAAAGA 1300

RESULT 14

US-10-380-731-119
Sequence 119, Application US/10380731
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-114
CURRENT APPLICATION NUMBER: US/10/380,731
CURRENT FILING DATE: 2003-03-14
PRIOR APPLICATION NUMBER: 09/659,671
PRIOR FILING DATE: 2000-09-11
NUMBER OF SEQ ID NOS: 888
SOFTWARE: Custom
SEQ ID NO 119
LENGTH: 2130
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (191)..(1426)
US-10-380-731-119

Query Match 100.0%; Score 90; DB 8; Length 2130;
Best Local Similarity 100.0%; Pred. No. 1.8e-22;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCGATGGCAGCGACATATAAATCC 60
Db 1211 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCGATGGCAGCGACATATAAATCC 12710
OY 61 AATCCCACTTGAACCACTGAGAGAAAAGA 90
Db 1271 AATCCCACTTGAACCACTGAGAGAAAAGA 1300

RESULT 15

US-09-532-315B-9504
Sequence 9504, Application US/09532315B
GENERAL INFORMATION:
APPLICANT: Sellhamey, Jeffrey J.
APPLICANT: Deleagane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullshy, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES
FILE REFERENCE: PD-1002 CIP
CURRENT APPLICATION NUMBER: US/09/532,315B
CURRENT FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 07/916,491
PRIOR FILING DATE: 1992-07-17
PRIOR APPLICATION NUMBER: 07/977,780
PRIOR FILING DATE: 1992-11-19
PRIOR APPLICATION NUMBER: 08/100,523
PRIOR FILING DATE: 1993-08-03
PRIOR APPLICATION NUMBER: 09/008,119
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/196,364
PRIOR FILING DATE: 1994-02-14
PRIOR APPLICATION NUMBER: 08/282,991
PRIOR FILING DATE: 1994-07-28
PRIOR APPLICATION NUMBER: 08/438,571
PRIOR FILING DATE: 1995-05-10
PRIOR APPLICATION NUMBER: 08/179,873
PRIOR FILING DATE: 1994-01-11

PRIOR APPLICATION NUMBER: 08/504,732
PRIOR FILING DATE: 1995-07-20
PRIOR APPLICATION NUMBER: 08/137,951
PRIOR FILING DATE: 1993-10-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 42212
SOFTWARE: PERL Program
SEQ ID NO 9504
LENGTH: 287
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
OTHER INFORMATION: Incyte ID No: hu00221292
FEATURE:
NAME/KEY: unsure
LOCATION: 107, 113, 120-121, 174, 259, 279, 282
OTHER INFORMATION: a, t, c, g, or other
US-09-532-315B-9504

Query Match 83.8%; Score 75.4; DB 6; Length 287;
Best Local Similarity 96.7%; Pred. No. 2.3e-17;
Matches 87; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 1 TGTGACCTCTGTGATCGATCATCATTTGGGATCGC-GAATGGCAGCGACATATAAATC 59
Db 178 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCGAATGGCAGCGACATATAAATC 237
OY 60 CAAATCCCACTTGAACCACTGAGAGAAAAG 89
Db 238 CAAATCCCACTTGAACCACTGAGAGAAAAG 267

Search completed: April 21, 2003, 18:41:28
Job time : 52.7649 secs

US-10-342-887-1692

Query Match
Best Local Similarity 100.0%; Score 90; DB 9; Length 1749;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGACGACATAAATCC 60
DB 826 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGACGACATAAATCC 885

QY 61 AAATCCCACTTGACCAACTGAAGAAAGA 90
DB 886 AAATCCCACTTGACCAACTGAAGAAAGA 915

RESULT 10

US-10-170-235-40554

Sequence 40554, Application US/10170235
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig
TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
FILE REFERENCE: C1001380
CURRENT APPLICATION NUMBER: US/10/170,235
CURRENT FILING DATE: 2003-03-17
NUMBER OF SEQ ID NOS: 42514
SEQ ID NO 40554
LENGTH: 2119
TYPE: DNA
ORGANISM: HUMAN
US-10-170-235-40554

Query Match
Best Local Similarity 100.0%; Score 90; DB 8; Length 2119;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGACGACATAAATCC 60
DB 1213 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGACGACATAAATCC 1272

QY 61 AAATCCCACTTGACCAACTGAAGAAAGA 90
DB 1273 AAATCCCACTTGACCAACTGAAGAAAGA 1302

RESULT 11

US-60-453-135-275

Sequence 275, Application US/60453135
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: IAKOUBOVA, Olga
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001456
CURRENT APPLICATION NUMBER: US/60/453,135
CURRENT FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 275
LENGTH: 2119
TYPE: DNA
ORGANISM: Homo sapiens
US-60-453-135-275

Query Match
Best Local Similarity 100.0%; Score 90; DB 11; Length 2119;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGACGACATAAATCC 60
DB 1213 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGACGACATAAATCC 1272

QY 61 AAATCCCACTTGACCAACTGAAGAAAGA 90

DB 1273 AAATCCCACTTGACCAACTGAAGAAAGA 1302

RESULT 12

US-60-453-050-275

Sequence 275, Application US/60453050
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: LIKE, May
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: STEMOSIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001457
CURRENT APPLICATION NUMBER: US/60/453,050
CURRENT FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 275
LENGTH: 2119
TYPE: DNA
ORGANISM: Homo sapiens
US-60-453-050-275

Query Match
Best Local Similarity 100.0%; Score 90; DB 11; Length 2119;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGACGACATAAATCC 60
DB 1213 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGACGACATAAATCC 1272

QY 61 AAATCCCACTTGACCAACTGAAGAAAGA 90
DB 1273 AAATCCCACTTGACCAACTGAAGAAAGA 1302

RESULT 13

US-09-659-671A-134

Sequence 134, Application US/09659671A
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Zhou, Ping
APPLICANT: Asundi, Vinod
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyang
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wehtman, Tom
APPLICANT: Dimanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 794
CURRENT APPLICATION NUMBER: US/09/659,671A
CURRENT FILING DATE: 2000-09-11
NUMBER OF SEQ ID NOS: 475
SOFTWARE: PL-FL-genes Version 2.0
SEQ ID NO 134
LENGTH: 2130
TYPE: DNA
ORGANISM: Homo sapiens
US-09-659-671A-134

Query Match
Best Local Similarity 100.0%; Score 90; DB 6; Length 2130;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGACGACATAAATCC 60

```

; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 276
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-050-276

Query Match
Best Local Similarity 100.0%; Score 90; DB 11; Length 1271;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAATCC 60
DB 365 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAATCC 424

QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
DB 425 AATCCCACTTGAACCACTGAAGAAAAGA 454

RESULT 6
US-10-170-235-41123
; Sequence 41123, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 41123
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-41123

Query Match
Best Local Similarity 100.0%; Score 90; DB 8; Length 1719;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAATCC 60
DB 813 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAATCC 872

QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
DB 873 AATCCCACTTGAACCACTGAAGAAAAGA 902

RESULT 7
US-60-453-135-274
; Sequence 274, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-274

Query Match
Best Local Similarity 100.0%; Score 90; DB 11; Length 1719;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAATCC 60
DB 813 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAATCC 872

QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
DB 873 AATCCCACTTGAACCACTGAAGAAAAGA 902

RESULT 8
US-60-453-050-274
; Sequence 274, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-050-274

Query Match
Best Local Similarity 100.0%; Score 90; DB 11; Length 1719;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAATCC 60
DB 813 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAATCC 872

QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
DB 873 AATCCCACTTGAACCACTGAAGAAAAGA 902

RESULT 9
US-10-342-887-1692
; Sequence 1692, Application US/10342887
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yundong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1692
; LENGTH: 1749
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1692
```

NAME/KEY: unsure
LOCATION: 228, 241
OTHER INFORMATION: a, t, c, g, or other
US-09-532-315B-9510

Query Match 100.0%; Score 90; DB 6; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.1e-22;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATCATTTGGGATCGCGAATGGCGACGACATATAATCC 60
DB 34 TGTGACCTCTGTGATGCAATCATCATTTGGGATCGCGAATGGCGACGACATATAATCC 93
QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
DB 94 AATCCCACTTGAACCACTGAAGAAAAGA 123

RESULT 2

US-09-532-315B-9519
Sequence 9519, Application US/09532315B
GENERAL INFORMATION:
APPLICANT: Sellhammer, Jeffrey J.
APPLICANT: Deleese, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Sluave, Laura L.
APPLICANT: Mullaby, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES
FILE REFERENCE: PD-1002 CIP
CURRENT APPLICATION NUMBER: US/09/532.315B
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 07/916,491
PRIOR FILING DATE: 1992-07-17
PRIOR APPLICATION NUMBER: 07/977,780
PRIOR FILING DATE: 1992-11-19
PRIOR APPLICATION NUMBER: 08/100,523
PRIOR FILING DATE: 1993-08-03
PRIOR APPLICATION NUMBER: 09/008,119
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/196,364
PRIOR FILING DATE: 1994-02-14
PRIOR APPLICATION NUMBER: 08/282,991
PRIOR FILING DATE: 1994-07-28
PRIOR APPLICATION NUMBER: 08/438,571
PRIOR FILING DATE: 1995-05-10
PRIOR APPLICATION NUMBER: 08/179,873
PRIOR FILING DATE: 1994-01-11
PRIOR APPLICATION NUMBER: 08/504,732
PRIOR FILING DATE: 1995-07-20
PRIOR APPLICATION NUMBER: 08/137,951
PRIOR FILING DATE: 1993-10-14
Remainder of prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 42212
SOFTWARE: PERL Program
SEQ ID NO 9519
LENGTH: 289
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: hu00815059
US-09-532-315B-9519

Query Match 100.0%; Score 90; DB 6; Length 289;
Best Local Similarity 100.0%; Pred. No. 1.1e-22;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATCATTTGGGATCGCGAATGGCGACGACATATAATCC 60
DB 84 TGTGACCTCTGTGATGCAATCATCATTTGGGATCGCGAATGGCGACGACATATAATCC 143
QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90

DB 144 AATCCCACTTGAACCACTGAAGAAAAGA 173

RESULT 3
US-10-170-235-41124
Sequence 41124, Application US/10170235
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig
TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HU
FILE REFERENCE: CL001380
CURRENT APPLICATION NUMBER: US/10/170,235
PRIOR FILING DATE: 2003-03-17
NUMBER OF SEQ ID NOS: 42514
SEQ ID NO 41124
LENGTH: 1271
TYPE: DNA
ORGANISM: HUMAN
US-10-170-235-41124

Query Match 100.0%; Score 90; DB 8; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1.6e-22;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATCATTTGGGATCGCGAATGGCGACGACATATAATCC 60
DB 365 TGTGACCTCTGTGATGCAATCATCATTTGGGATCGCGAATGGCGACGACATATAATCC 424
QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
DB 425 AATCCCACTTGAACCACTGAAGAAAAGA 454

RESULT 4

US-60-453-135-276
Sequence 276, Application US/60453135
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: IAKOBOVA, Olga
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CL001456
CURRENT APPLICATION NUMBER: US/60/453,135
PRIOR FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 276
LENGTH: 1271
TYPE: DNA
ORGANISM: Homo sapiens
US-60-453-135-276

Query Match 100.0%; Score 90; DB 11; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1.6e-22;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATCATTTGGGATCGCGAATGGCGACGACATATAATCC 60
DB 365 TGTGACCTCTGTGATGCAATCATCATTTGGGATCGCGAATGGCGACGACATATAATCC 424
QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
DB 425 AATCCCACTTGAACCACTGAAGAAAAGA 454

RESULT 5

US-60-453-050-276
Sequence 276, Application US/60453050
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: LUKE, May
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:37:18 ; Search time 50.7649 Seconds
(without alignments)
7947.624 Million cell updates/sec

Title: US-09-513-151-3_COPY_1121_1210

Perfect score: 90
Sequence: 1 TGTGACCTGTGATCGAAT.....TGACCACTGAAGAAGA 90

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 5930009 seqs, 2241446263 residues

Total number of hits satisfying chosen parameters: 11860018

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents, NA, New:*

1: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US12_NEW_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US13_NEW_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US14_NEW_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US15_NEW_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US16_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	263	6	US-09-532-315B-9510 Sequence 9510, Ap
2	90	100.0	289	6	US-09-532-315B-9519 Sequence 9519, Ap
3	90	100.0	1271	8	US-10-170-235-41124 Sequence 41124, A
4	90	100.0	1271	11	US-60-453-135-276 Sequence 276, App
5	90	100.0	1271	11	US-60-453-050-276 Sequence 276, App
6	90	100.0	1719	8	US-10-170-235-41123 Sequence 41123, A
7	90	100.0	1719	11	US-60-453-135-274 Sequence 274, App
8	90	100.0	1719	11	US-60-453-050-274 Sequence 274, App
9	90	100.0	1749	9	US-10-342-887-1692 Sequence 1692, Ap
10	90	100.0	2119	8	US-10-170-235-40554 Sequence 40554, A
11	90	100.0	2119	11	US-60-453-135-275 Sequence 275, App
12	90	100.0	2119	11	US-60-453-050-275 Sequence 275, App
13	90	100.0	2130	6	US-09-659-671A-134 Sequence 134, App
14	90	100.0	2130	6	US-10-380-731-119 Sequence 119, App
15	90	100.0	287	6	US-09-532-315B-9504 Sequence 9504, Ap
16	74	82.2	2103	9	US-10-144-771-2119 Sequence 2119, Ap
17	74	80.4	301	7	US-09-539-806B-46638 Sequence 46638, A
18	52.8	58.7	434	6	US-09-912-293-54990 Sequence 54990, A
19	47.2	52.4	248	9	US-10-349-781-405285 Sequence 405285, A
20	47	52.2	246	9	US-10-349-781-405285 Sequence 405285, A
21	29.8	33.1	241	6	US-09-532-315B-9512 Sequence 9512, Ap
22	28.8	32.0	1037984	7	US-09-947-911-259 Sequence 259, App

23	28	31.1	358	11	US-60-141-233-66580 Sequence 66580, A
24	27.6	30.7	259	7	US-09-837-604A-36790 Sequence 36790, A
25	27.6	30.7	441	7	US-09-837-604A-80102 Sequence 80102, A
26	27.6	30.7	458	7	US-09-837-604A-18601 Sequence 18601, A
27	27.6	30.7	462	7	US-09-837-604A-52531 Sequence 52531, A
28	27.6	30.7	519	7	US-09-837-604A-35887 Sequence 35887, A
29	27.6	30.7	593	7	US-09-837-604A-54885 Sequence 54885, A
30	27.6	30.7	605	7	US-09-837-604A-53708 Sequence 53708, A
31	27.6	30.7	22345	6	US-09-849-002-681 Sequence 681, App
32	27.6	30.7	22345	6	US-09-849-002-746 Sequence 746, App
33	27.6	30.7	22347	6	US-09-849-002-745 Sequence 745, App
34	26.8	29.8	599	8	US-10-115-678-136 Sequence 136, App
35	26.8	29.8	599	8	US-10-112-944-136 Sequence 136, App
36	26.6	29.6	601	7	US-09-849-016-162940 Sequence 162940, A
37	26.6	29.6	42118	7	US-09-849-016-16297 Sequence 16297, A
38	26.4	29.3	3474	9	US-10-144-771-5708 Sequence 5708, Ap
39	26.4	29.3	3687	9	US-10-144-771-21447 Sequence 21447, A
40	26.2	29.1	396	11	US-60-141-233-45678 Sequence 45678, A
41	26.2	29.1	807	6	US-09-134-000C-1231 Sequence 1231, Ap
42	26.2	29.1	807	6	US-09-134-000C-1231 Sequence 1231, Ap
43	26.2	29.1	24837	7	US-09-849-016-16177 Sequence 16177, A
44	26	28.9	200	6	US-09-532-315B-9517 Sequence 9517, Ap
45	26	28.9	201	11	US-60-453-135-65757 Sequence 65757, A

ALIGNMENTS

RESULT 1
US-09-532-315B-9510
Sequence 9510, Application US/09532315B
GENERAL INFORMATION:
APPLICANT: Sellhammer, Jeffrey J.
APPLICANT: Deleage, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullaly, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES
FILE REFERENCE: PD-1002 CIP
CURRENT APPLICATION NUMBER: US/09/532, 315B
CURRENT FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 07/916,491
PRIOR FILING DATE: 1992-07-17
PRIOR APPLICATION NUMBER: 07/977,780
PRIOR FILING DATE: 1992-11-19
PRIOR APPLICATION NUMBER: 08/100,523
PRIOR FILING DATE: 1993-08-03
PRIOR APPLICATION NUMBER: 09/008,119
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/196,364
PRIOR FILING DATE: 1994-02-14
PRIOR APPLICATION NUMBER: 08/282,991
PRIOR FILING DATE: 1994-07-28
PRIOR APPLICATION NUMBER: 08/438,571
PRIOR FILING DATE: 1995-05-10
PRIOR APPLICATION NUMBER: 08/179,873
PRIOR FILING DATE: 1994-01-11
PRIOR APPLICATION NUMBER: 08/504,732
PRIOR FILING DATE: 1995-07-20
PRIOR APPLICATION NUMBER: 08/137,951
PRIOR FILING DATE: 1993-10-14
Remaining prior application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 42212
SOFTWARE: PERL Program
SEQ ID NO 9510
LENGTH: 263
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: hu00416615
FEATURE:

Alignment Scores:

Pred. No.:	2,046-18	Length:	311
Score:	334.50	Matches:	93
Percent Similarity:	51.52%	Conservative:	60
Best Local Similarity:	31.31%	Mismatches:	103
Query Match:	9.36%	Indels:	41
DB:	2	Gaps:	11

US-09-513-151-3 (1-2041) x F64046 (1-311)

```
QY 86 GTAAGTATTCCTCGGGCCGCGCCAGCCGCAAAATCCAGCGTGGCTTGACGTAGCCAG 145
   ::::::::::: ||| ::::::::::: |||
Db 6 IlePheLeuMetGlyProThrAlaSerGlyThrSpleuAlaIleGlnLeuArgSer 25
   ::::::::::: ||| ::::::::::: |||
QY 146 CGGCTGGCGCGGTGATGTCAGCGGTGACTTCATGCGAGGTCTATGAGGCGCTAGACATC 205
   ::::::::::: ||| ::::::::::: |||
Db 26 GlnLeuProValGluValIleSerValAspSerAlaLeuIleTyrLysGlyMetAspIle 45
   ::::::::::: ||| ::::::::::: |||
QY 206 ATGACCAACAAGTTTCTGCCCAAGACAGAGAAATCGCCGCGCACACATGATCAGCTT 265
   ||| ||| ||| ::::::::::: |||
Db 46 GlyThrAlaLysProSerLysGlnGlnLeuAlaProHisArgLeuIleAspIle 65
   ::::::::::: ||| ::::::::::: |||
QY 266 GTGATCCCTCTGTGACCAATTAACAGGTGGTGGACTTCAGAAATAGCAACTGCTCTG 325
   ::::::::::: ||| ::::::::::: |||
Db 66 LeuAspPro---SerGluSerTyrSerAlaMetAsnPheArgAspAlaLeuArgGlu 84
   ::::::::::: ||| ::::::::::: |||
QY 326 ATTGAAGATATATTGGCCGACGCAAAATTCCTATTGTTGGGAGGACCAATTATTAC 385
   ::::::::::: ||| ::::::::::: |||
Db 85 MetAlaAspIleThrAlaGlnGlyLysIleProLeuLeuValGlyGlyThrMetLeuTyr 104
   ::::::::::: ||| ::::::::::: |||
QY 386 ATTGAATCTCTGCTGTGGAAGTCTGTGCAATACCAAGCCCCGAGAGATGGCAGCTGAG 445
   ::::::::::: ||| ::::::::::: |||
Db 105 TyrLysAlaLeu-----IleGlnGlyLeuSerProLeuProSerAlaAspGlu 120
   ::::::::::: ||| ::::::::::: |||
QY 446 AAAGTGAATGACGCAAAAGTGAAGCTTGAA-----AAGGAGATGCTCTGTA 493
   ::::::::::: ||| ::::::::::: |||
Db 121 AsnIle-----ArgAlaGlnLeuGlnGlnLysAlaIleGlnGlnGlyTyrAlaAla 137
   ::::::::::: ||| ::::::::::: |||
QY 494 CTTACACAAGCGCTAGCCAGGTGGACCCGAAATGCGTGGCAAGCTGCATCCACATGAC 553
   ||| ||| ||| ::::::::::: |||
Db 138 LeuHisThrGlnLeuAlaLysIleAspProIleSerAlaAlaArgIleAsnProSerAsp 157
   ::::::::::: ||| ::::::::::: |||
QY 554 AAAGCGAAAGTGGCCAGAGCTTCGAGTTTGAAGAACAGGAATCTCATAGTGA 613
   ::::::::::: ||| ::::::::::: |||
Db 158 SerGlnArgIleAsnArgAlaLeuGlnValPheTyrIleThrGlyLysSerLeuThrGlu 177
   ::::::::::: ||| ::::::::::: |||
QY 614 TTCTCCATCGTCACATACGAGAGAGTGTGTGTCCTCGAGGTCTCTGAAGTTC 673
   ||| ||| ||| ::::::::::: |||
Db 178 Leu-----ThrGlnGlnLysGlnLysAlaLeu-----ProTyrAspPhe 190
   ::::::::::: ||| ::::::::::: |||
QY 674 TCTAACCTTGATCCTTGGCTTCATGCTGACCGACGAGTCTTAGATGAGCGCTTGAT 733
   ||| ||| ||| ::::::::::: |||
Db 191 ValGlnPheAlaIle-----AlaProGlnAspArgHisValLeuHisGlnArgIleGlu 208
   ::::::::::: ||| ::::::::::: |||
QY 734 AAGAGGTGATGATGATGCTGCTGCTGGGCTGTGGAGAACTAGAGATTTTCACAGA 793
   ::::::::::: ||| ::::::::::: |||
Db 209 GlnArgPheHisLysMetIleGlnLeuGlyPheGlnAlaGlnValGlnLysLeuTyrAla 228
   ::::::::::: ||| ::::::::::: |||
QY 794 CGCTATATATCAGAAATGTTTCGAAATAGCCAGAGCTATCAACATGTATCTTCCAA 853
   ||| ||| ||| ::::::::::: |||
Db 229 Arg---GlyAspLeuAsnIleAsnLeuProSer-----IleArg 240
   ::::::::::: ||| ::::::::::: |||
QY 854 TCAATTTGGCTTCAAGAAATTTACAGACTACCTG----- 886
   ::::::::::: ||| ::::::::::: |||
Db 241 CysValGlnTyrArgGlnMetTyrPglutTyrLeuGlnGlnLysPyrAlaTyrGlnGlnMet 260
   ::::::::::: ||| ::::::::::: |||
QY 887 ATCAGTAGGGAATAATACACTGAGAGCTAGTAACAGCTTCTAAGAAA 937
   ||| ||| ||| ::::::::::: |||
Db 261 IlePheArgGlnIleCys-----AlaThrArgGlnLeuAlaLysArg 274
   ::::::::::: ||| ::::::::::: |||
```

Search completed: April 21, 2003, 18:51:09
Job time : 56.8883 secs

```

OY 614 TTTCATCGTCACATACGAGAGAGGTGGTGGTCCCTTGAGAGTCTCTGAGTTTC 673
DB 187 LysVal---GlnGlnPheHrlLysn----- 194
OY 674 TCTACCCCTTGCTCTTGCTTCATGCTGCACAGCAGTCTTAGATGAGCGCTTGAT 733
DB 195 TyrAspThrIleLeuLeuGlyIleGlnMetSerArgLysThrLeuTyrSerIleLysn 214
OY 734 AAGAGGGTGGATGACATGCTGCTGCTGGGCTCTTGAGAGAACTAAGAGATTTCACAGA 793
DB 215 LysArgValAspIleMetLeuAspHisGlyLeuPheArgGluVal----- 229
OY 794 CGCTTAATCAAGAAATGTCGGAATAATACCAAGACATACATGATCTTCCAA 853
DB 230 -----GlnGlnLeuValGlnGlnGlyTyrGlnSerCysGlnSer-----MetGln 244
OY 854 TCAATGGCTTCAGAA 871
DB 245 AlaIleGlyTyrLysGlu 250

RESULT 14
T05359
Hypothetical protein F22K18.150 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 26-Aug-1999
C:Accession: T05569
R:Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hohnsels, J.; Mewes, H.W.; Mayer, K.F.
Submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15419
A:Accession: T05569
A:Molecule type: DNA
A:Residues: 1-318 <BEY>
A:Cross-references: EMBL:AL035356
A:Experimental source: cultivar Columbia; BAC clone F22K18
A:Genetics:
A:Map position: 4
A>Note: F22K18.150
C:Superfamily: delta(2)-isopentenylpyrophosphate transferase

Alignment Scores:
Pred. No.: 1.08e-18 Length: 318
Score: 338.00 Matches: 97
Percent Similarity: 42.23% Conservative: 58
Best Local Similarity: 26.43% Mismatches: 90
Query Match: 9.45% Indels: 122
Gaps: 10

US-09-513-151-3 (1-2041) x T05569 (1-318)
OY 83 CTGTAGATCTCTGGGGCCGACCGGACCGCAATCCAGCCTGGCTGACGTAGGC 142
DB 7 MetValIleMetGlyAlaThrIleSerGlyLysSerLeuSerValAspLeuAla 26
OY 143 CAGCGGCTCGGGGTGATGCTGACCGCTGACCTGATGAGAGGTATGAGCGCTAGAC 202
DB 27 LeuHisPheLysAlaGlnIleLeuSerSerLysMetGlnPheThrAspIleLys 46
OY 203 ATCATCACCAACAGAGTTTCTGCGCCACAGAGAGAGATCTGCGCGCACCATGATCAGC 262
DB 47 IleThrThrAsnGlnSerThrIleGlnLysPheArgGlyValAlaProHisIleLeuGly 66
OY 263 TTGTGGATCTCTGTGATACCAATACACAGTGGAGCTTCAGAAATAGACCAACTGCT 322
DB 67 GluLeuAsnProGluAlaGlyGluValThrAlaGlnPheArgValMetAlaAlaGlu 86
OY 323 CTGATTGAAGATATATTGGCCGAGACAATAATCTTATGTTGGGAGGAACCAATTAT 382
DB 87 AlaIleSerGluIleThrGlnArgLysLysLeuProIleLeuAlaGlyLysSerAsnSer 106
OY 383 TACATGAATCTGCTCTGGAAGATTCTGTCAATACCAAGCCCAAGAGATGGGCACCT 442
DB 107 TyrIleHisAlaLeuLeuAlaLys----- 114

```

```

OY 443 GAGAAAGTATGATGACCCGAAAGTGGAGCTTGAAGAGAGATGCTTGTACTTGACAA 502
DB 114 ----- 114
OY 503 CCGCTAAGCAGGTGGACCCAGAAATGGCTGCCAAGTGCATGCATGACAAAGCAAA 562
DB 115 -----SerTyrAspProGlns-----TyrProPheSerAspHisLys 127
OY 563 GTGGCCAGAGCTTCAAGATTGTAAGAAACAGAAATCTCATAGTAATTTCTCCAT 622
DB 127 ----- 127
OY 623 CGTCAACATACGAGAAAGGTGGTGGTCCCTTGAGAGTCTGTGAATTTCTTAACCT 682
DB 128 -----GlySerIleCysSerGlnLeuLysTyr---AspLys 138
OY 683 TGCATCTTTGGCTTCATGCTGACAGGACAGTGTAGATGAGCGCTGATTAAGAGTys 742
DB 139 CysPheIleThrIleAspValAspIleSerValLeuPheGlnIleAlaGlnPheHisArg 158
OY 743 GATGACATGCTGCTGCTGCTGGGCTCTTGAGAGAACTAAGATTTTCACAGACGTATAT 802
DB 159 AspLeuMetMetLysSerGlyMetPheGlnGlnIleAlaGlnPheHisArg----- 175
OY 803 CAGAAAGATGTTTTCGGAATAATAGCCAGACTATCACTATGATGTTCCAAATTCAGC 862
DB 176 -----SerLysLysAlaProLysGluProLysGlyIleThrLysAlaIleGly 191
OY 863 TTCAGGAATTTACAGATACCTG-----ATCAGTACGAGGAAATATGC 886
DB 192 ValGlnGluPheAspAspTyrLeuLysMetTyrLysTrpAspAsnAspMetAspLysTrp 211
OY 887 -----ATCAGTACGAGGAAATATGC 904
DB 212 AspProMetArgLysGluAlaIleThrGlnLysAlaValAlaGlnAlaLysGlnThrPhe 231
OY 905 ACACGTGAGACTAGTAACCAAGCTTTAAAG-----AAAGACACTGGTCCCAATT 952
DB 232 GlnLeu---ThrLysAspGlnIleThrLysIleAsnLysLeuArgAsnAlaGlyTrpAsp 250
OY 953 GTCCTCCCTGTC-----TATGAC 970
DB 251 IleLysLysValAspAlaIleThrAlaSerPheArgGlnAlaIleArgAlaLysGlnGly 270
OY 971 TTAGAGTATGATGATGCTGAG-----TGGAGAGAGTCTGTTGAACTGCTCTT 1024
DB 271 GluGlyValAlaGlnMetGlnArgLysIleThrAsnLysGluValLeuGlnProCysVal 290
OY 1025 GAATCTGTCGAAGTTTATC 1045
DB 291 LysIleValArgSerHisLeu 297

RESULT 15
F64046
delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.-) - Haemophilus influenzae
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
C:Accession: F64046
R:Flisshmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavar,
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fulmerman, J.L.; Geoghagen, N.S.I
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vent
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: F64046
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-311 <TRIG>
A:Cross-references: GB:U32692; GB:U42023; NID:91573013; PIDN:AC21746.1; PID:915730;
C:Superfamily: delta(2)-isopentenylpyrophosphate transferase
C:Keywords: nucleotidyltransferase; tRNA modification

```

C:Genetics:
A:Gene: miaA
C:Superfamily: delta(2)-isopentenylpyrophosphate transferase

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	3,356-19	339	98	63	115	40
Percent Similarity:	344.50					
Best Local Similarity:	51.11%					
Query Match:	31.11%					
DB:	2					9

US-09-513-151-3 (1-2041) x A71473 (1-339)

```

QY 41 CCTGTGGGCGATGGGCTCAGAGGCGCTGCAACGACCTACTCTTGTAGTATCTCGG 100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19 ProGlnLysSerPheSerLysMetPheLysArgThr-----ValIleLeuLeuAlaGly 36
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 101 GCCACGGGCAACCGGCAATCCAGCTGGCGTTGCAGTAGAGCGGCTGGCGGTGAG 160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37 ProThrIlySerGlyLysThrAlaValSerLeuLysLeuAlaProLeuValAspGly 56
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 161 ATCTGACGCGTACCTCATGACGCTATGAGCGCTAGACATCATCACCACAGGTT 220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 57 IleIleSerValAspSerMetGlnValTyrGlnGlyMetAspIleGlyThrAlaLysVal 76
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 221 TCTGCCACAGACAGACAGATCTGCCGACACATGATCAGCTTTGTGATCCTTGTG 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 77 SerLeuThrAspArgLysGlnValAlaProHisIleLeuLeuAsp---ValCysHisValGln 95
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 281 ACCAATTAACAGTGGGAGTCTGAGAAATAGACAACTGCTGATGATGAGATATATT 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 GluSerPheAsnAlaValAspPheTyrThrHisAlaValGlnAlaCysGlnAspIleLeu 115
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 341 GCCCGAGACAAATTCCTATTGTGTGGGAGAACCAATTAATTAATTAATTCCTGCTC 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 116 SerArgAsnLysValProIleLeuValGlyGlyThrGlyPheTyrPheHisThrPheLeu 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 401 TGGAAAGTCTTTCATATACCAAGCCCGAGAGATGGGACACTGAGAAAGTATTACCG 460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 -----SerGlyProProSerGlyProSerProAspPheValLeuArgLys 150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 461 AAATGGAGCTTGAA---AAGGAGATGGTCT---GTATCTCAACAAAGGCTAGCCAG 514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151 GlnLeuThrLeuGlnAlaGlnGlnLysGlyIleSerAlaLeuTyrGlnGlnLeuLys 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 515 GTGACACCAAGAAATGGCTGCCAAGCTGCATCCATGACAAACGCAAGTGGCCAGAGC 574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 LeuAspProValTyrAlaAlaThrIleThrLysHisAspLysAsnLysIleIleArgAla 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 575 TTGCAAGTTTGAAGAAACAGAAATCTCTCATAGTGAATTTCTCCATCGTCACATACG 634
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 LeuGlnIleIleArgLysThrGlySerLysValSerSerTyrAlaTrpGlnSerThrVal 210
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 635 GAAGAGGTGGTGGTCCCTGGAGGCTCTCTGAGATCTCTCAACCTTGACCTCTTGG 694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 AsnGlnSerLys-----GluTyrHisCysArgArgTyr 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 695 ---CTTCATGCTGACACAGCAGATTCTAGATGAGCCTTGATAGAGGGTGGATGACATG 751
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 222 LeuLeuSerProAspProGlnLeuLeuArgHisAsnIleLeuGlnAlaArgCysAspGln 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 752 CTGCTGCTGGCTCTTGGAGACTAAGAGATTTTTCACAGACGCTATTAATAGAAAGT 811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 LeuGlnGlnGlnLysLeuLeuAspLysValGlnAlaLeuLeuAlaGlyTyrLeuSerLys 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 812 GTTTCGGAAGAAATAGCCAGGACTATCAACATGATGATCTCCATCAATGGCTCAGAGA 871
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 262 SerSerAlaSer-----ArgAlaIleGlyTyrArgGln 272
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 872 TTTCACGAGTACTG-----ATCAGTACGAGGAGAAA 901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 273 TrpIleGlnPheLeuAspLeuGlnLysSerProProAspLeuPheGlnIleThrLysGlnLys 292

```

QY 902 TGCACACTGGAGACTAGTAACCA-GCTTTAAGAAAGACCTG 945
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 293 PheIleThrAsnThrTrpArgTyrThrLysLysGlnArgThrTrp 307

RESULT 13

B89905
 hypothetical protein miaA [imported] - staphylococcus aureus (strain N315)
 C:Species: Staphylococcus aureus
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: B89905
 R:Kuroda, M.; Ohla, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.;
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu,
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.
 Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant staphylococcus aureus.
 A:Reference number: A89758; M01D:21311952; PMID:11418146

A:Accession: B89905

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-311 <KUR>

A:Cross-references: GB:BA000018; PID:g13701103; PIDN:BA842398.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: miaA

C:Superfamily: delta(2)-isopentenylpyrophosphate transferase

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	8,966-19	311	86	62	92	26
Percent Similarity:	339.00					
Best Local Similarity:	55.64%					
Query Match:	32.33%					
DB:	2					7

US-09-513-151-3 (1-2041) x B89905 (1-311)

```

QY 83 CTGTAGAGATTCCTGGGCGGCGGACCGGCAACATCCAGCTGGGCTGACGTAGGC 142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8 IleValValIleValGlyProThrAlaSerLysThrGlnLeuSerIleGlnLeuAla 27
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 143 CAGCGGCTCGGCGGTGAGATGTCAGCGCTGATCCATGACGCTGATGAGGCTTAGAC 202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28 LysArgIleAsnGlyGlnIleIleSerGlyAspSerMetGlnValTyrLysHisMetAsn 47
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 203 ATCATCCAAACAGGTTTGGCCCAAGACAGACAGAAATCTCCGCGACCAATGATCAGC 262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 48 IleGlyThrAlaLysValThrProGlnGlnMetAspGlyIleProHisHisLeuIleAsp 67
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 263 TTTGGATCGCTTGTGACCAATTAACACAGTGGGACCTCAGAAATAGACCAACTGCT 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68 IleLeuAsnProAspAspThr---PheSerAlaTyrGlnPheLysArgLeuAlaGlnAsp 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 323 CTGATTGAAGATATATTGCCCAGACAAATTCCTATTGTGTGGGAGAACCAATTAT 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 87 LeuIleThrAspIleThrAsnArgGlyLysValProIleIleAlaGlyThrGlyLeu 106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 383 TACATTGAATCTCTGCTGCGAAAGTCTTGTCAATACCAAGCCCGACAGATGGGCACT 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 107 TyrIleGlnSerLeuIleTyrAsnTyrGlnLeuGlnAspGlnThrValThrProAlaGln 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 443 GAGAAAGTATGATGACCGGAAAGTGAG---CTTGAAGAGAGATGCTTGTACTTCAC 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127 LeuSerIleValLysGlnLysLeuSerAlaLeuGlnHisLeuAspAsnGlnGlnLeuHis 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 500 AAACGCTTAAGCCAGGTGAGCCAGCAAAATGGCTGCCAAGCTGCATCCACATGACAAAGC 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 147 AspTyrLeuAlaGlnPheAspAlaValSerAlaGlnAsnIleHisProAsnAsnArgGln 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 560 AAAGTGGCCAGGAGCTTGCACGTTTGAAGAAACAGA-----ATCTTCATAGTGA 613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 ArgValLeuArgAlaIleGlnTyrTyrLeuLysThrLysLysLeuLeuSerAsnArgLys 186

```


Db 103 TyrIleAsnSerIleuIleCysAsnTyrAspPheThcGlyAlaTyrIleCysAspGluAlaTyr 122
 QY 440 ACTGAGAAAGTATTGACCGAAAAGTGAGCTTGAAGAGAGAGTGTCTT---GTACTT 496
 Db 123 ArgGluSerLeu-----GlnAlaIleAlaIleAspGlyLysGluTyrLeu 138
 QY 497 CACAAAGGCTTAAGCCAGGTGACCCAGAAATGGCTGCCAAGCTGCATGCATGACAA 556
 Db 139 HisGluLysLeuLysAsnIleAspSerTyrLysIleLysLeuTyrProAsnAspLeu 158
 QY 557 CCGAAGTGGCCGAGAGCTTGCAGTTTGAAGAAGAGATCTCATGTGATTT 616
 Db 159 LysArgValIleArgAlaIleGluIleValTyrLysIleThcGlyLysThrIleSerGluLeu 178
 QY 617 CTCATCGTCACATACGAGAGAGAGGTGTGCTCCCTTGGAGAGTCTTGAAGTTCTCT 676
 Db 179 -----AsnSerAsnValAspLeuTyrAspIleProTyr----- 189
 QY 677 AACCTTGCATCTTGGCTTCACTGCTGACACGAGCTTCTAGATAGCGCTTGATAG 736
 Db 190 AsnIleHisTyrPheIleLeuAsnMetAspArgGlnLysLeuTyrGluArgIleAsnLeu 209
 QY 737 AGGCTGATGACATGCTTCTGCTGCTGCTGCTGAGAGAA-----CTAGAGATTTT 787
 Db 210 ArgValAspIleMetLeuArgAsnGlnLysValAspGluValIleLysLeuArgAspMet 229
 QY 788 CACAGAGCTTAATACAGAGATGTTTCGAAAATAGCCAGACTATACATGATGATC 847
 Db 230 GlyTyrAsnSerAsnMetClnSerMet----- 238
 QY 848 TTCATATCATTTGGCTTCAAGATTTTCAGAGTACCTGATACGAGGAAATGACACA 907
 Db 239 ---LysGlyIleGlyTyrLysGlnIleuSerTyrLeu-----GlnGlyCysIleThr 255
 QY 908 CTGAGACTAGTAACACGCTCTTAAGAAAGA 940
 Db 256 LeuGluGluAlaValGlnIleuIleLysGly 266
 RESULT 11
 C72366
 tRNA delta-2-isopentenylpyrophosphate transferase - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: C72366
 R:Neelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A:Reference number: A72200; MUID:99287316; PMID:10360571
 A:Accession: C72366
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-305 <ARN>
 A:Cross-references: GB:AE001728; GB:AE000512; NID:94981027; PIDN:AA035610.1; PID:9498103
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM0525
 C:Superfamily: delta(2)-isopentenylpyrophosphate transferase
 Alignment Scores:
 Pred. No.: 1,72e-19 Length: 305
 Score: 348.00 Matches: 86
 Percent Similarity: 53.66% Conservative: 68
 Best Local Similarity: 29.97% Mismatches: 105
 Query Match: 9.73% Indels: 28
 DB: 2 Gaps: 8
 US-09-513-151-3 (1-2041) x C72366 (1-305)
 QY 77 CTACCTCTGTAGTATTCTCGGCGACGACCGACGCAATCAACGCTGCGTTCAG 136
 Db 1 MetLysIleAlaIleValGlyLysProThrAlaValGlyLysThrAspIleMetIleGlu 20

QY 137 CTAGCCACAGGCTCGCGGTGATGATGTCAGCGCTGACTCCATGACGCTATGAAAGC 196
 Db 21 ValCysGluGluIleIleGlyAlaGluIleIleSerMetAspSerArgIleIleTyrArgTyr 40
 QY 197 CTAGACATCATCACCAAGAGTTTGTGCCCAAGACAGAGAGATTCGCGGACACAG 256
 Db 41 MetAspIleGlyThrAlaLysProThrProGluGlnArgLysArgValLeuHisIleMet 60
 QY 257 ATCAGCTTGTGATGATCTTGTGACCAATTTACACAGTGGTGTGACTTCGAATATAGCA 316
 Db 61 IleAspIleIleSerPro---AspGluTyrTyrAsnAlaPheMetTyrArgLysAspSer 79
 QY 317 ACTGCTCTGATGAGATATATTTGCGCGACCAAAATTCATATGTTGTGGAGAGAC 376
 Db 80 LeuArgAlaMetGluAspAlaLeuArgAlaGlyLysIleIleProValTyrAlaGlyLysThr 99
 QY 377 AATTATTCATGATGATCTGCTCTGGAAGATTTCTGTCAATACCAAGCCCAAGAGATG 436
 Db 100 GlyLeuTyrAlaAspAlaLeuVal---ArgGlyIlePheGluGluGlyAlaProAlaAspGlu 118
 QY 437 GGCACAGAGAACTGATTCACCGAAAGTGGAGCTTGAAGAGAGATGCTTGTACTT 496
 Db 119 AsnIleArgLysGluLeuArg-----GluLeuGluArgArgGluProGluIleLeu 135
 QY 497 CACAAAGGCTTAAGCCAGAGTGGACCCAGAAATGGCTGCCACAGCTGATCCATGACAA 556
 Db 136 ArgLysMetLeuGluGluLeuAsnProGluAlaAlaThrArgIleHisProAsnAspLeu 155
 QY 557 CGCAAGTGGCCGAGAGCTTGCAGATTTTGAAGAAAGAGAAATCTCATAGTGAATTT 616
 Db 156 LysArgThrIleArgAlaLeuGluIleValTyrMetLysThrGlyArgArgIleSerGluLeu 175
 QY 617 CTCATATCATATACATACGAGAGTGTGCTGCCCTTGGAGAGTCTTGTGATGATCTCT 676
 Db 176 -----GlnLysGluAlaLysGlyAspAsp-----ArgPhe--- 185
 QY 677 AACCTTGCATCTTGGCTTCACTGCTGACACGAGCTTCTAGATAGCGCTTGATAG 736
 Db 186 -----PheIleIleValLeuThrArgGluArgTyrGluLeuTyrGluArgIleAsnLys 203
 QY 737 AGGCTGATGACATGCTTCTGCTGCTGCTTGGAGAGAACTAAGATTTTCAACAGCC 796
 Db 204 ArgValAspLysMetIleGluMetGlyLeuValAspGluValLysArgLeuGluMet 223
 QY 797 TTAATCAACAAGATGTTCCGAAAATACCAAGACTATCAACATGATCTTCCAATCA 856
 Db 224 GlyTyrSerLysAspLeuAsnSer-----MetLysThr 234
 QY 857 ATGGCTTCAGAGATTTTCACAGATACCTGATCAGTACAGAGAAATGCACACTGAGACT 916
 Db 235 IleGlyTyrLysGluValIleAspTyrLeu-----GlnGlyLysTyrAspPheAspLys 252
 QY 917 ACTAACACGCTTCTTAAGAA 937
 Db 253 MetValHisLeuIleLysArg 259
 RESULT 12
 A71473
 probable tRNA pyrophosphate transferase - Chlamydia trachomatis (serotype D, strain
 C:Species: Chlamydia trachomatis
 C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
 C:Accession: A71473
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mltc
 Science 282, 754-759, 1998
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia
 A:Reference number: A71570; MUID:9900809; PMID:9784136
 A:Accession: A71473
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-339 <ARN>
 A:Cross-references: GB:AE001349; GB:AE001273; NID:93329226; PIDN:AA068361.1; PID:93
 A:Experimental source: serotype D, strain UW-3/Cx


```

|||||  |||  |||:|||||  |||||  |||||:|||||
47 AsplIegLYThAlAlaYsleTherGlUGlnGluMetGlUGlyValProHisHisLeuIle 66
OY 260 ACCTTGGATGATCTCTGTGTGACCAATACAGTGGTGGACTTCAGAAATAGACCACT 319
Db 67 AsplIeuaNsPro---GlnAspSerPheSerThAlAspTylGlnSerLeuValArg 85
OY 320 GCTGATGATGAAATATTTTCCCGAGACAAATTCATATGTTGTGTGGAGAACCAAT 379
Db 86 AsnLYsleSerGlUleAlaAsnArgGlyLYsLeuProMetCileAspLYsIleThrGly 105
OY 380 TATTACATTGAATCTCTGCTGTGAAAGTTCTTGCATATACCAAGCCCGAGAGATGGGC 439
Db 106 LeuTYrIleGlnSerGluLeuTYrAspTYrThrPheThr-----GluGluAla 121
OY 440 ACTGAGAAAGTATGATGACCGAAAGTGGAGCTT-----GAAAGAGAGATGGCTTGTGA 493
Db 122 AsnaSProValPheArgGluSerMetGlnMetAlaAlaGluArgGluLYsAlaAspPhe 141
OY 494 CTTCAGAAAGCCGTTAAGCGAGTGGACCCGAAATGGCCGCAAGCTGCATCCACATGAC 553
Db 142 LeuHISAlaLYsLeuAlaAlaAlaAspProGluAlaAlaAlaAlaIleHisProAsna 161
OY 554 AAACGCAAGTGGCCAGAGCTTGCAGATTGTAAGAAACAGCAATCTCATAGTGA 613
Db 162 ThrArgArgValIleArgAlaLeuGluIleLeuHISThrSerGlyLYsThrMetSerGln 181
OY 614 TTTCTCATGCTGCACATACAGGAAGGTGGTGTCCCTCCCTGGAGGCTCTGAAATTC 673
Db 182 HISLeuLYsGluGlnLYsArgGlu-----LeuLeu 191
OY 674 TCTAACCCCTGCATCCTTTGGCTTCATGCTGACAGAGCACTTACATAGCGCTTGAT 733
Db 192 TYrAsnaAlaValLeuIleLYsLeuThrMetAspArgSPThrLeuTYrIleArgIleAsn 211
OY 734 AAGAGGCTGATGATGATGCTGTGCTGGCTGAGAGCACTAAGAAATTTTCACAGA 793
Db 212 GlnArgValAspLeuMetMetGlnSerGlyLeuLeuProGluVal-----LYs 227
OY 794 GCGTATATACAGAAAGATGTTTCGAAATAAGCCAGACTATACATGATCTTCCAA 853
Db 228 ArgLeuTYrAspLYsAsnaVal-----ArgAspCysGlnSer-----IleGln 241
OY 854 TCAATTGGCTTCAAGAAATTTACAGAGTACCTGCATCAGCAGAGGAAATGCACACTGGAG 913
Db 242 AlaIleGlyTYrLYsGlnLeuTYrAlaTYrPhe-----AspGlyPheValIleLeuSer 259
OY 914 ACTAGTAAACGACTTCTAAAGAA 937
Db 260 AspAlaValGluGlnLeuLYsGln 267

RESULT 8
AF1236
tRNA isopentenylpyrophosphate transferase homolog miaa [imported] - Listeria monocytogenes
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AF1236
R:Glasner, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.;
J. Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Me
Ox, C.; Schlueter, T.; Simoes, N.; Tlherre, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1236
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-305 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC93972.1; PID:g16410710; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: miaa

```

```

C:Superfamily: delta(2)-isopentenylpyrophosphate transferase
Alignment Scores:
Pred. No.: 2e-23 Length: 305
Score: 397.50 Matches: 98
Percent Similarity: 54.51% Conservative: 59
Best Local Similarity: 34.03% Mismatches: 100
Query Match: 11.12% Indels: 31
Gaps: 7
DB: 2

US-09-513-151-3 (1-2041) x AF1236 (1-305)
OY 77 CTACCTGTGATGATGATTTCTGGGGCCAGCGGCAATTCACAGCTGGCTTGAG 136
Db 4 IleProValIleValIleValAlaGlyProThrAlaValGlyLYsThrSerLeuSerIleGlu 23
OY 137 CTAGCCAGCGGCTGGCGGTGAGATGCTGACGCGTGCATCCATCCAGTCTATGAAAGC 196
Db 24 LeuAlaLYsLYsLeuAspLYsGluIleIleSerGlyAspSerMetGlnValTYrArgLY 43
OY 197 CTAGACATCATACCAACAGGTTCTGCCCCAGAGCAGACAGATCTCCCGCACCATG 256
Db 44 LeuAspIleGlyThrAlaLYsIleThrProGluGluMetAspGluIleLYsHisTYrLeu 63
OY 257 ATCAGCTTGTGATCCTCTTGTGACCAATTACAGAGTGGCTGACTTCAGAAATAGAGA 316
Db 64 IleAspValThrAspPro---SerGluProPheThrAlaAlaLYsPheIleThrGluThr 82
OY 317 ACTGCTGTGATGAGATATATTTCGCCGAGCAAAATTCATATTGTTGGAGAACCC 376
Db 83 ArgLYsTrpIleGluThrIleHisGlnAlaGlyLYsLeuProIleIleValGlyTYr 102
OY 377 AATTATTCATGAAATCTCTCTGGAAGATCTTGTCAATACCAAGCCCGAGAGATG 436
Db 103 GlyLeuTYrIleGlnSerValPheTYrAspTYrAspPheGlyAsnValSerGluAspLYs 122
OY 437 GGCCTAGAGAAAGTATTCAGCCGAAAGTGAGCTTGAAGAGAGATGCTTGTACT 496
Db 123 Ala-----TYrArgAlaGluLeuGlnGluLeuAsnLYsThrLeu 136
OY 497 CACAAAGCCCTAAGCAGGTGAGCCAGAAATGGCTGCCACGCTCATCCATGACAA 556
Db 137 TrpGlnMetLeuGlnGlnAspProGluSerAlaAlaGlnIleHisGlnAsnLYs 156
OY 557 GCGAAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGAAATCTTCATATGAATT 616
Db 157 ArgArgValIleArgAlaLeuGlnValMetHISLeuThrGlyLYsPheProPheSerGluTYr 176
OY 617 ---CTCCATGCTGCACATCAGGAAGAGTGGTGGTCCCTCGAGAGCTCTGAAATTC 673
Db 177 GlnValIleHisAsnValLeuAsnAspThr----- 185
OY 674 TCTAACCCCTGCATCCTTTGGCTTCATGCTGACAGCAAGCACTTACATAGCGCTTGAT 733
Db 186 TYrLYsProLeuPheLeuGluLYsLeuAspArgAlaLeuLeuTYrGluArgIleAsn 205
OY 734 AAGAGGCTGATGATGATGCTGTGCTGGCTTGGAGAGCACTAAGAAATTTTCACAGA 793
Db 206 GlnArgValAspLeuMetPheGlnGluGlyLeuValThrGluAlaLYs-----LYs 222
OY 794 GCGTATATACAGAAAGATGTTTCGAAATAAGCCAGAGTATACATGATCTTCCAA 853
Db 223 LeuTYrAspGlnHisLeuValAspValProAla-----ValLYs 235
OY 854 TCAATTGGCTTCAAGAAATTTACAGAGTACCTGCATCAGCAGGAAATGCACACTGGAG 913
Db 236 GlyIleGlyTYrLYsGlnLeuPheProTYrPhe-----GluGlyLYsSerSerLeuGlu 253
OY 914 ACTAGTAAACGACTTCTAAAGAA 937
Db 254 GlnAlaLYsGluLeuIleGlnLYs 261

RESULT 9

```

RESULT 6

C RNA isopentenylpyrophosphate transferase miaA [imported] - Bacillus halodurans (strain C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: F83945
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirata
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: F83945
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-314 <STO>
A:CROSS-references: GB:P001515; GB:BA000004; NID:g10174886; PIDN:BA06085.1; GSPDB:GN000
C:Experimental source: strain C-125
C:Genetics:
A:Gene: miaA
C:Superfamily: delta(2)-isopentenylpyrophosphate transferase

Alignment Scores:

Pred. NO.:	5,67e-25	Length:	314
Score:	417.00	Matches:	102
Percent Similarity:	56.27%	Conservative:	64
Best Local Similarity:	54.58%	Mismatches:	85
Query Match:	11.66%	Indels:	44
DB:	2	Gaps:	10

US-09-513-151-3 (1-2041) x F83945 (1-314)

OY 83 CTGTAGTGAATTCCTGGGCGCCACGGGCACGGCAATCCAGCTGCCCTTGACCTAGGC 142
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5 leuValAlaIleValAGlProThrAlaValAlaLysTrhThServalMetLeuLa 24

OY 143 CAGGCGCTCGGGGAGATCGTAGCGCTGACTCATGACGAGTTATGAAGGCTTAAC 202
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25 LysArgLeuAnGLyGIuValIIleSerGIyAspSerMetGIuValTYrArgGIyMetAsp 44

OY 203 ATCATCACCAACAAGGTTCTGCCCAGAAGCAGAAATCTGCCGCCACCACATGATCAGC 262
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 45 IlegIThrAlaTyIleThrAlaGlugluMetAspGIyAlaProHISileuIleAsp 64

OY 263 TTTGTGATCCCTCTGTATACCAATTACAGAGTGGTGGACTTCAGAAATAGAACACTGCT 322
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65 IleLysAspPro--SerGIuSerPheSerValAlaAspPheGlnAspLeuLatHrPro 83

OY 323 CTGATTGAAGAAATATTTGGCCGAGACAAATAATCTTATTGTGTGGAGAAACCAATAT 382
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 84 LeuIleThrGIuIIleHisGIuArgGIyArgLeuProPheLeuValGIelYThrGIyLeu 103

OY 383 TACATTGAATCTCTGCTGTGAAGATT--CTGTCAATTACCACGCCCCAGAGATWGAGC 439
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 104 TyrValAsnAlaValIIleHisGIuPheAsnLeuGIyAspRIleArgAlaAspLu----- 121

OY 440 ACTGGAAGATGATGACCGAAAAGTGGAGCTGAA-----AAGAGAGATGCTT 490
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 -----AspTyrrArgHISgluLeuGlnAlaPheValAsnSerTyrgIyVal 136

OY 491 ---GTACTTCACAAAGCGCTAAGCCAGAGTGGACCCAGAAATAGCTGCCACGTGATCCA 547
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137 GluAlaLeuHIAspLyLeuSerTyrlAspProLYsAlaAlaAlaAlaIleHisPro 156

OY 548 CATGACAAACGCAAGTGGCCAGAGACTTGCAAGTTTTTGAAGAAACAGAAATCTCTCAT 607
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 157 AsnAsnTyrrArgArgValIIleArgAlaLeuGIuIleIleTyrlLeuThryLysThrVal 176

OY 608 AGTGAATTTTCATCGCAACATPACGAAGAAGGTGGTGGCTTGGAGAGTCTCTG 667
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 177 Thglu-----GlnlaaargHISglUGluGluThrProseProryr----- 190

Qy 668 AAGTTCCTAAACCTTCGATCCCTTTGGCTTCATGCTGACGACGACATTTCTAGATGAGCC 727

Db 191 -----AsnLeuValMetIleGlyLeuThrMetLeuArgAspAlaLeuArgAspArg 207

Qy 728 TTGGATAAGAGCGTGATGATGACATGCTGTGCTGGGCTCTGGAGGACATAAGATTTT 787

Db 208 ILeSnatgArgValAspGlnMetValGlnGluGlyLeuIleAspGlnAlaIlyslsLeu 227

Qy 788 CACAGACGCTTAATCATGAGAAGATGTTTCGGAAATATAGCCAGGACTATCAACATGSTATC 847

Db 228 -----TyrAspArgIlyle 232

Qy 848 -----TTCCATCATATGCGCTTCAAGGAATTTACAGAGTACCTGATCACT 892

Db 233 ArgAspCysGlnSerValGlnAlaIleGlyTyrIlyslsMetIlyAspTyrIleu----- 250

Qy 893 GAGGAAATATGCACATCGAGACACTAGTAACACGCTTATMAAGAA 937

Db 251 AspGlyAsnValThrLeuGluGlnAlaIleAspThrLeuIlyslsArg 265

RESULT 7

669657

tRNA isopentenylpyrophosphate transferase miaA - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence-revision 05-Dec-1997 #extLchange 20-Jun-2000

C:Accession: G69657

R:Kunitz, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, C.; Broth, S.; Brouillett, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, A.; Ehrlich, S.D.; Emmertson, P.T.; Estlin, K.D.; Erlington, J.; Fabret, C.; F Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizel, Iech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Koester, P.; Konligstein, G.; Krogh, S.; Kumano, M.; Kunita, K.; Lapidus, A.; Laubert, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, Y. M.; Ogawa, K.; Ogikawa, A.; Oudeg, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T. A:Authors: Schleich, S.; Schroeter, R.; Scofield, F.; Sekiguchi, J.; Sekowaka Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yaguchi, K.; Yata, K.; A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A. A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus A:Reference number: A69580; MUID:R8044033; PMID:9384377

A:Accession: G69657

A:Status: Preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-314 <RUND>

A:Cross-references: GB:299113; GB:AL009126; NID:g7634090; PIDN:CAB13617.1; PTC:Genetics:

A:Experimental source: strain 168

C:Genetics:

A:Superfamily: delta(2)-isopentenylpyrophosphate transferase

Alignment Scores:

Pred. No.: 6.71e-24 Length: 314

Score: 403.50 Matches: 97

Percent Similarity: 55.56% Conservative: 63

Best Local Similarity: 33.68% Mismatches: 29

Query Match: 11.29% Indels: 99

DB: 2 Gaps: 8

US-09-513-151-3 (1-2041) x G69657 (1-314)

Qy 80 CCTCTTGAAGATTCTTCGGGCGGACGACGACCGCAATCCAGCGTGGCGTTCAGCTA 139

Db 7 ProValAlaIleLeuValGlyProThrAlaValGlyIlyslsThrAsnLeuSerIleGlnLeu 26

Qy 140 GGCCAGCGCGTGGCGGCGTGAGATGCTGACGCGCTGACTCCATGACGAGCTATGAAGCCCTA 199

Db 27 AAllyssSerLeuAsnAlaGlnIleIleSerGlyAspSerMetGlnIleIlyslsGlyMet 46

Qy 200 GACATCATCACCAACAGGTTTCTGCCCAAGACGAGAAATCTGCCGACACCATGATC 259

OY 989 -----TCGACGTGGAGAGACGTCCTGTCTGAACCTGCTTGAATGCGCA 1036
 DB 371 SerlysserGluGluSerTrpAsnAlaGlnValValysProAlaSerGluIleLeuArg 390
 OY 1037 AGTTTCATCCAG-----GGCCACAAGCCTACAGCCATCCATAAAGATG 1081
 DB 391 CysPheLeuGluThrGlnThrGluSerGlyArgAspProThrSerGlyLys----- 407
 OY 1082 CCATACAAATGAGCTGAGACACAGAGAAGATTATCACTGTGTGACCTGTG--GATGCA 1138
 DB 408 -----SerIleGluArgAspLeuThrPheGlnThrGlnValValysGluAlaCysGlyAsnLys 425
 OY 1139 ATCATCATGGGGGATGGCGGATGGCGACGACATAAATCCATCCACTTG----- 1192
 DB 426 IleLeuArgGlyArgHisGluTrpGlnHisHisLysGlnGlyArgThrHisArgLysArg 445
 OY 1193 -----ACCAACTGAAGAAAGAGAGAGAGATTGAGCTGACATGCTGTC 1234
 DB 446 ThrThrArgHisLysAsnSerGlnThrTyrrLysAsnArgGluValGlnGluAlaGluVal 465
 OY 1235 AAC 1237
 DB 466 Asn 466

RESULT 5
 hypobacterial protein At2g27760 (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_rev10 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: F84676
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Umayam, L.; Tallon, L.;
 Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: F84676
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-461 <STO>
 A:Cross-references: GB:AE002093; NID:93860256; PID:NAC73024.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g27760
 A:Map position: 2

Alignment Scores:
 Pred. No.: 3,2e-29 Length: 461
 Score: 471.00 Matches: 128
 Percent Similarity: 48.96% Conservative: 107
 Best Local Similarity: 26.67% Mismatches: 137
 Query Match: 13.17% Indels: 108
 DB: 2 Gaps: 17

US-09-513-151-3 (1-2041) x F84676 (1-461)

OY 65 CTGCAACGACCTACCTCTGTAGTATTCGGGGCAGCGGACCGGCAATCCACAG 124
 DB 1 MetLysLysValAlaValAlaValIleMetGlyProThrGlySerGlyLysSerLys 20
 OY 125 CTGGGTTGACAGCTAGCGGCGGCTCGGCGGTGAGATGTCAGGCTGACTCCATGCA 184
 DB 21 LeuAlaValAspLeuAlaSerHisPheProValGluIleIleAsnAlaAspLametin 40
 OY 185 GTCTATGAGGCTAGACATCACCAACAAGTTTCTGCCACAGACAGAGATCGC 244
 DB 41 IleTrpSerGlyLeuAspValLeuThrAsnLysValThrValAspGlnLysGlyVal 60
 OY 245 CGGACACATGATCAGCTTTGTGATCCTCTTGACCAATTCACAGAGTGCTGACTTC 304
 DB 61 ProHisHisLeuLeuGlyThrValSerSerAspMet-----GluPheThrAlaArgSph 79
 OY 305 AGAATATAGACACTGCTCTGATTAAGATATATTGCCCGAGACAAATTCCTATGCTT 364

DB 80 ArgAspPheThrValProLeuIleGluIleValaSerArgAsnHisIleProValLeu 99
 OY 365 GTGGAGGAGACCAATTTATCATGATGATCTGCTGGAAGTTCTTGCAATACCAAG 424
 DB 100 ValGlyGlyThrHisIleTyrIleGlnAlaValAlaSerLysPheLeuAsnAspAspAla 119
 OY 425 CCCCGAGAGTGGGACACTGAGAA-----GTATGACCGGAAA 463
 DB 120 AlaGluAsp-----ThrGluGluCysCysAlaAspValAlaSerValValAspIleAsp 137
 OY 464 GTGGAGCTTGA-----AAGAGGATGCTCTTGACTACATCAAAAGCCTTAAGC 511
 DB 138 MetValValGluSerValPheGlyArgAspAspLeuSerHisGlyTyrGluLeuLys 157
 OY 512 CAGGTGACCCAGAAATGCTGCCAAGCTGCATCCATGACACAAAGCAAGTGGCCAGG 571
 DB 158 GluLeuAspProValAlaAlaAsnArgIleHisProAsnAsnHisArgIleLeuGln 177
 OY 572 AGCTTGCAAGTTTGAAGAAACAGAAATCTCTCATGATTTCTCATCGTACAT 631
 DB 178 TyrLeuSerLeuHisAlaSerArgGlyValLeuProSerLysLeuTyrGlnGlyThr 197
 OY 632 ACGGAA-----GAAGGTGGTGCCTCCCTGGAGCTGCT 664
 DB 198 AlaGluValAspSerPhePheLeuThrSerArgGlnAsnTrpGlyCysIleAsnAlaSer 217
 OY 665 CTGAGTTCTCTAACCTTGCACTCTTGCTTGGCTTCAAGCTGACAGCAAGTTCTAGATGAG 724
 DB 218 ---ArgPhe---AspTyrCysIleLeuIleCysMetAspIleGluThrAlaValLeuAspArg 235
 OY 725 CGCTTGATTAAGAGGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 784
 DB 236 TyrValGlnGluArgValAlaAspLametinValAspAlaGlyLeuLeuAspGluValLysAsp 255
 OY 785 TTTCACAGACGCTATATACAGAAAGATTTTGGAAATAGCCAGACATCAACATGCT 844
 DB 256 IleTyrLys-----ProGlyAlaAspTyrThrArgGly 266
 OY 845 ATCTTCATCAATGATGCTTCAAGAAATTCACAGACTCTG-----ATCACTAGG 895
 DB 267 LeuArgGlnSerIleGlyValAlaArgLysPheGlnAspPheLeuLysIleHisLeuSerGlu 286
 OY 896 -----GGAAATGACACA---CTGGAGACTACTAACACAGCTTCTAAAGACAGACT 943
 DB 287 ThrCysAlaGlyHisLeuThrSerLeuSerAsnAspAspLysValIleMetLysGluAsnLeu 306
 OY 944 GGTCCCAT----- 952
 DB 307 ArgLysIleLeuAsnPheProLysAspAspLysLeuArgIleMetLeuGluGluAlaIle 326
 OY 952 ----- 952
 DB 327 AspArgValLysLeuAsnThrArgArgLeuLeuAlaArgGlnLysArgValSerArg 346
 OY 953 GTCCCCCTGTCTATGCTTACAGGTAATCTGATGTC----- 988
 DB 347 LeuGluThrValPheGlyTrpAsnIleHisTyrIleAspAlaThrGluTyrIleLeuSer 366
 OY 989 -----TCGAATGGGAGAGACTGTCTTGAACCTGCTCTGGAATAGCTGCAAGCT 1039
 DB 367 LysSerGluGluSerTrpAsnAlaGlnValAlaLysProAlaSerGluIleIleLeuArgCys 386
 OY 1040 TTCATCCAG-----GGCACAAGCCTACAGCCATCCATAAAGATGCA 1084
 DB 387 PheLeuGluThrGlnThrGluSerGlyArgAspProThrSerGlyLys----- 402
 OY 1085 TACAATGAGCTGAGACACAGAGATTATCACTGTGTGACCTGTG--GATGCAATC 1141
 DB 403 ---SerIleGluArgAspLeuThrPheGlnThrGlnValValCysGluAlaCysGlyAsnLysIle 421
 OY 1142 ATCATTTGGGAGATGGGATGGGACGACGACATAAATCCATCCACTTG----- 1192
 DB 422 LeuArgGlyArgHisGluTrpGlnHisHisLysGlnGlyArgThrHisArgLysArgThr 441

```

Db 210 AspSerLeuValLeuMetProArgLeuAspLysArgValAspLysMetLeuSerHisGly 229
QY 764 CTCCTGGAGGAAGTAAGATTTTCACAGACCTAATACAGAGATGTTCCGGAAT 823
Db 230 LeuValAspGluIleLys-----SerMetLysSerLeuAlaGluSer 243
QY 824 -----AGCAGAGATCATCAATGATCTTCATCAATGATGCTTCAGAAATTT 874
Db 244 GluLysPheSerProAspPheThrArgGlyIleTyrPheLysIleLysGluPhe 263
QY 875 CACGAGTACCTG-----ATCAGTACGAGAAATGCACACTGGAG----- 913
Db 264 MetProTyrPheGluAlaProSerAspIleValPheAsnAspCys---LeuGluArgMet 282
QY 914 -----ACTAGTACCAAGCTTTAAAGAA----- 937
Db 283 LysValSerThrArgIleTyrAlaLysSerGlnLysTyrIleGlnSerArgPheLeu 302
QY 938 -----GGACCTGGTCCATTGTCCCTCCCTGCTATGGC 970
Db 303 PrometCysLeuAlaGlnLysLeuSerProSerSerIleLeu-----PheSerThr 320
QY 971 TTAGAGATCTGATGCTCGAAGTGGAGAGTCTGTTCTTGAACCTGCTCTTGAATC 1030
Db 321 ThrAsnThrThrAspLeuAsnAsnTyrGluGlnVal---GluLysAlaCysArgVal 339
QY 1031 GTGCAAAAGTTTCATCCAGGCGCCACAAAGCTACAGCCATCCATTAAGATCCATCAAT 1090
Db 340 PheGlnTyrPhePheTyrAsnGlyAspAlaIleAla---ProSerAlaAspAspGlnHis 358
QY 1091 GAAGCTGAGACAGACAGAGAGTAT----- 1114
Db 359 AlaPheGluLysAlaArgAspTyrLeuSerIleMetAsnGlyArgGlnSerGlnLys 378
QY 1115 ---CACCTGTGACCTCTGT---GATCGA-----ATCATCATTTGGGAGAT 1153
Db 379 LysPheValCysGluLysLeuAspLysArgGlyAspProPheThrValIleGlyGlu 398
QY 1154 CCGGAATGGGACGACGACATTAATCCAAATCCCACTTGACCACTGAAGAAAGAGA 1213
Db 399 AspAlaPheAsnValHisIleLysSerArgLysHisLysThrThrValArgArgLys 418

RESULT 4
T52061
tRNA Isopentenyltransferase (EC 2.5.1.8) [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C:Accession: T52061
R:Golovko, A.; Hjaltn, G.
submitted to the EMBL Data Library, November 1998
A:Description: A tRNA Isopentenyl transferase from Arabidopsis thaliana.
A:Reference number: Z25928
A:Accession: T52061
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-466 <GOL>
A:Cross-references: EMBL:AF109376; PIDN:AAF00582.1
C:Keywords: transferase

Alignment Scores:
Pred. No.: 5,72e-31 Length: 466
Score: 493.00 Matches: 133
Percent Similarity: 50.108 Conservative: 108
Best Local Similarity: 27.658 Mismatches: 138
Query Match: 13.798 Indels: 102
DB: 2 Gaps: 18

US-09-513-151-3 (1-2041) x T52061 (1-466)
QY 41 CCTGTGGGCGAGTGGCTCAGAGGC-----CTGCACGCGACCTTCTTGTGATGATT 94
Db 6 ProSerAsnGluLysIleGluGlnLysMetLysLysAlaLysValValIle 25

```

```

QY 95 CTCGGGGCCAGCGGACCGCAAAATCCAGCTGGCGTTGACCTAGCGCCGCTCGGC 154
Db 26 MetGlyProThrGlySerCylLysSerLysLeuAlaValAspLeuAlaSerHisPhePro 45
QY 155 GGTGAGATGTCAGCGGCTGACTCCATGACGAGTCTATGAAGGCTAGACATCAACAC 214
Db 46 ValGluIleIleAsnAlaAspAlaMetGlnIleTyrSerGlyLeuAspValLeuThrAsn 65
QY 215 AAGGTTTCCGCCAGAGAGAGAAATCCCGCGCCACACATGATACAGCTTGTGATCCT 274
Db 66 LysValThrValAspGluGlnLysGlyValProHisHisLeuMetGlyThrValSerSer 85
QY 275 CTGTGACCAATTACACAGCTGCTGACCTCAGAAATAGACAACCTCTGTATGAAGT 334
Db 86 AspMet---GluPheThrAlaArgAspPheAlaGAspPheThrValProLeuIleGlu 104
QY 335 AATATTGCCGAGACAAATTCCTATGTTGTTGGAGAGAACCAATATTATTAATGAATCT 394
Db 105 IleValSerArgAsnHisIleProValIleGlyGlyThrHisTyrTyrIleGlnAla 124
QY 395 CTGCTGTGAAAGTTCTTGTCAATACCAAGCCCGAGAGATGGCAGCTAGAAA----- 448
Db 125 ValValSerLysPheLeuLeuAspAspAlaAlaGluAsp---ThrGluLysCys 142
QY 449 -----CTGATTCACGAGAAAGTGGACCTTGAA-----AAGAG 481
Db 143 AlaAspValAlaSerValAlaAspGlnAspMetValValGluSerValPheGlyArgAsp 162
QY 482 GATGCTCTTGTACTTCACAAAGCCCTTACGCAAGTGGAGCCAGAAATGGCTGCCAGCTG 541
Db 163 AspLeuSerHisGlyTyrIleLeuLeuLysGluLeuAspProValAlaAlaAsnArgIle 182
QY 542 CATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGATTGTTGAACAAACAGAAATC 601
Db 183 HisProAsnAsnHisArgLysIleAsnGlnTyrLeuSerLeuHisAlaSerArgGlyVal 202
QY 602 TTCATATGAAATTTCTCCATGCTCAACATGAGAGAAAGTGGTCCCTGGAGAGT 661
Db 203 LeuProSerLys---LeuTyrGlnGlyLysThrAlaGlnAsnTyrGlyLysIleAsnAla 221
QY 662 CCTCTGAAGTCTCTAACCCTTGACATCCCTTGGCTTCATGTCAGCAGCAGCTTCTAGAT 721
Db 222 Ser---ArgPhe---AspTyrCysLeuIleCysMetAspAlaGluThrAlaValLeuAsp 239
QY 722 GAGCCCTGGATTAAGAGGTGATGATGATGCTGCTGCTGGCTTGGAGAGACATAA 781
Db 240 ArgTyrValGluGlnArgValAspAlaMetValAspAlaGlyLeuLeuAspGluValTyr 259
QY 782 GATTTCACAGACGCTATATCAGAGAAATGTTTGGAAATATGACGAGCATCAACAT 841
Db 260 AspIleTyrLys-----ProGlyAlaAspTyrThrArg 270
QY 842 GGTATCTCCATCAATTTGGCTTCAAGAAATTCAGAGTACCTG-----ATCACT 892
Db 271 GlyLeuArgGlnSerIleGlyValArgGluPheGluAspPheLeuLysIleHisLeuSer 290
QY 893 GAG-----GGAATATGCACA---CTGGAGACATGTAACAGCTTCTTAAGAAGAGA 940
Db 291 GluThrCysAlaGluHisIleThrSerLeuSerAsnAspAspValMetLysGluAsn 310
QY 941 CTTGTGCCCAT----- 952
Db 311 LeuArgLysIleLeuAsnPheProLysAspAspLysLeuArgIleMetLeuGluGluAla 330
QY 952 ----- 952
Db 331 IleAspArgValLysLeuAsnThrArgArgLeuLeuArgGlnLysArgArgValSer 350
QY 953 ---GTCCCCCTGCTATAGGCTTAGAGATCTGATGATC----- 988
Db 351 ArgLeuGluThrValPheGlyTyrPheAsnIleHisTyrIleAspAlaThrGluTyrIleLeu 370

```

```

OY 308 AATGAGCAACTGCTGATT----- 328
Db 82 GluValThrLeuAspLeuIleValLeuAsnSerProLeuphegLeuAspProAsn 101
OY 329 -----GAGATATATTGCCCCGAGACAAATTCCTATTGTTGGGAGAAC 376
Db 102 PheHisAsnhecinlysIleArgAlaArgSerLysIleProValIleValGlyThr 121
OY 377 AATTATTACATGATCTGCTGTGAAA-----GTCTGTGCAATACCAACCCC 427
Db 122 ThrtYrGrlaGluSerValLeuYrGluAsnLeuIleGluThrAsnThrSerAsp 141
OY 428 CAGAGATGGGCACTGGAAAGTATGACCGAAAGTGGAGCTGAAAAGAGATGCT 487
Db 142 AspValAspSerLysSerAlaGlyThrSerSerGluSerSerGluAspThrGluGly 161
OY 488 CTT-----GTACTTCACAAAGCCCTAAGCCAGGTGGACCCAGAAATGGCTGCCAG 538
Db 162 IleSerAsnGlnGluLeuThrPaspGluLeuLysIleAspGluLysSerAlaLeuLeu 181
OY 539 CTGCATCCACATGACAAACGAAAGTGCCAGAGCTTGCAAGTTTGAAGAAACAGA 598
Db 182 LeuHisProAsnAsnArgYrArgValGlnArgAlaLeuGlnIlePheArgGluThrGly 201
OY 599 ATCTCATGATGATTTCTCCATCCGTCACATACGGAAGAAGGTGGTCCCTTGA 658
Db 202 IleArgLysSerGluLeuValGluLysGlnLysSerAspLys---ThrValAspLeuGly 220
OY 659 GGTCTCTAGATTTCTCTAATCCCTTGACCTTGCTGATGCTGACAGGAGCTTCTA 718
Db 221 GlyArgLeuArgPheAsnAsnSerLeuValIlePheMetAspAlaThrProGluValLeu 240
OY 719 GATGAGCGCTGGATTAAGAGGTGATGATGATCTGCTGCGCTTGGAGGAACATA 778
Db 241 GlnGlnArgLeuAspArgValAspLysMetIleLysLeuGlyLeuLysAsnGluLeu 260
OY 779 AGAGATTTTCAGAGCGCTTAATCAGAAAGATGTTCCGGAATAATGCCAGACTATCAA 838
Db 261 IleGluPhe-----TyrAsnGlnHis-----AlaGluYrIleAsnHisSerLys 275
OY 839 CATGATATCTTCAATCAATATGGCTTCAGGAATTTCCAGACTACTG----- 886
Db 276 TyrGlyValMetGlnCysIleGlyLeuLysGlnPheValProThrLeuAsnLeuAspPro 295
OY 887 -----ATCAGCTGAGGGAATGC-----ACA 907
Db 296 SerGluArgAspThrLeuAsnGlyAspLysLeuPheLysGlnLysCysAspValLys 315
OY 908 CTGGAGACT-----AGTAACGAGTTTCTAAG 934
Db 316 LeuHisThrArgGlnTyrAlaArgArgGlnArgGlyTrpYrArgSerAlaGlyLeuLys 335
OY 935 AAGAGCACTGGTCCC---ATTGTCCCCCTGTATGCTTAGAGATGATGATGCTGC 991
Db 336 ArgSerAspLysAspArgLysMetAlaSerThrLysMetLeuAspThrSerAsp----- 353
OY 992 AATGGAGAGCTGCTTCTTGAACCTGCTTGAATGCTGCAAACTTCATCCAGGC 1051
Db 354 -----LysTyrArgIleIleSerAspGlyMetAspIleValAspGlnThrMetAsnGly 371
OY 1052 -----CAGAGCGCTACAGCCACTCCAAATGAATGCCATACATGAAGTGAAGAC 1102
Db 372 IleAspLeuPheGlnAspIleSerThrAspThrAsn---ProIleLeuLysGlySerAsp 390
OY 1103 AAGAGAAGTTATCAGCTGTGTGACCTGTGTGATGCAATCTCATTTGGGAGTCCGAGATG 1162
Db 391 AlaAsnIleLeuLeuAsnCysGlnLysCysAsnIleSerMetThrGlyLysAspAsnTrp 410
OY 1163 GAGGCGACATAAATCCAAATCCATGGAACCAACTGAAGAAAGAGA 1213
Db 411 GlnLysHisIleAspArgLysLysHisLysHisIleAlaLysGlnLysLys 427

```

```

T38664
tRNA Isopentenyltransferase - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
R/Murphy, L., Harris, D., Wood, V., Rajandram, M.A., Barrell, B.G.
Submitted to the EMBL Data Library, August 1997
A/Accession number: Z21804
A/Reference number: T38664
A/Molecule type: DNA
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Residues: 1-434 <MUR>
A/Cross-references: EMBL:AL109739, NID:e1534774, PIDN:CA852278.1, GSPDB:GN00066, SPI
A/Experimental source: strain 972h-; cosmid c343
C/Genetics:
A/Genes: SPDB:SPAC343.15
A/Map position: 1

Alignment Scores:
Pred. No.: 8,46e-34 Length: 434
Score: 528.50 Matches: 141
Percent Similarity: 50.00% Conservative: 79
Best Local Similarity: 32.05% Mismatches: 133
Query Match: 14,788 Indels: 87
DB: 2 gaps: 19

US-09-513-151-3 (1-2041) x T38664 (1-434)
OY 80 CCGTGTAGATTTCTGCGGGCCGCGCAATCCAGCTGGCGTTCAGCTA 139
Db 4 ProLeuGlyValIleGlyThrGlyAlaGlyLysSerAlaValGlnLeu 23
OY 140 GCGCAGCGCTCGCGGTGAGATGTCAGCGCTACTCATGACAGTGTATGAAGCCTA 199
Db 24 AlaLysArgPheGlySerGlnValIleAsnAlaAspMetGlnIleYrArgGlyPhe 43
OY 200 GACATCTACACAAAGAGTTTCTGCCAAGACAGAGAAATCTCCGGCACCATATATC 259
Db 44 AspThrIleThrAsnLysIleThrValGlnGlnLysValHisAsnArgLeuMet 63
OY 260 AGCTTTGTGAGCTCTTGTGACCAATTCAGTGGTGGCTCAGAAATGAGCAACT 319
Db 64 SerPheLeuAsn---PheAspLysGlnTyrSerValProGluPheGluAlaGlyAspLysSer 82
OY 320 GCTGTGATGATATATTTGCCCGAGACAAATTCATTTGTTGGAGGAACCAAT 379
Db 83 ArgValIleAspGlnIleHisSerGlnGlyLysIleProIleValIleGlyThrHis 102
OY 380 TATTACATTGAATCTGCTGTGAAA-----GTCTGTGTC 415
Db 103 TyrTrpLeuGlnSerLeuLeuPheGlnAspThrThrLeuSerAlaIleAspLysLeuThr 122
OY 416 AAT-----ACCAAGCCCCAGAGAGATGGGCACTGGAAGATGATGACCGAAA 463
Db 123 AsnAspSerSerProSerLysProProHisProAspSerHis---IleLeuAsp----- 139
OY 464 GTGAGCTGAAAAGAGATGTCTGTGATCTTCACAAAGCGCTTAAGCCAGGTGAGCCA 523
Db 140 -----AspAspProSerAlaMetLeuSerYrLeuLysLysIleAspPro 154
OY 524 GAATGGCTGCCAAGCTGCAATCCACATGACAAACGCAAAAGTGGCCAGAGCTTCAAGTT 583
Db 155 ValMetAlaGlnGlnThrPheAspArgAspThrArgLysIleArgAspSerLeuGlnIle 174
OY 584 TTGAGAAGAACAGAAATCTCATATAGATTTCTCATGTCATGTAACATAGGAAGAGT 643
Db 175 TyrPheHisThrGluArgProProSerGlnIleYrSerGlnGlnLysMetLysSerSer 194
OY 644 GGTGTGCCCTGGAGAGTCTCTGAAGTCTTAACCCCTTGACATCCTTGGCTTCATGCT 703
Db 195 -----GlySerLysLeuArgTyrLysSer---LeuIlePheThrAlaPheAla 209
OY 704 GACCAGGCACTTCTAGATGAGCCGCTTGATTAAGAGGTGATGACATGCTTGGCTGGG 763

```

RESULT 3


```

Db 287 ThirgtrhrarggintyratlatyargglvallystrprrilelyslsmetleuilePro 306
OY 959 CCTGTCTATGGC-----TTAGAGTATCTATGTCTTCGAAAGTGGAGGACTCT 1006
Db 307 AspilelysglyaspiletyrLeuleuuaasplahsrpleusergintppasrPhasn 326
OY 1007 GTTGTGACCTGCTCTTGAANTGTCGCAAGTTTCATC-----CAG 1048
Db 327 AlaserglnatgalailealleleserAsnasphelleaserasnargproilleysgln 346
OY 1049 GGGCACAAGCCTTACAGCCACTTCATTAAGATGCCATTAATGCAAGCTGAGAACAGAGA 1108
Db 347 gluatrgalaprolysaaleuglunleuleuserlysglyglntrhrthmetlysls 366
OY 1109 -----ACTTTACACCTGTGTGACCTCTGT-----GATCGA 1138
Db 367 leuaspasprtrprrhrhistrYrthCysasnaValcysargasnaalaaapglylsasnaVal 386
OY 1139 ATCATCATTTGGGATGCGGAATGGGAGCGCACGCAATAAATCCAAATCCACTGGAACCA 1198
Db 387 ValalaileglicglyltsYrtrPrlgs1lehtlsleuglyserargarghnlsserAsn 406
OY 1199 CTGAAGAAAGACAGAGA 1216
Db 407 leuysargasnaThrarg 412

RESULT 2
T27538
hypothetical protein ZC395.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence-revision 15-Oct-1999 #text-change 15-Oct-1999
C:Accession: T27538
R:Connell, M.
submitted to the EMBL Data Library, August 1994
A:Description: The sequence of C. elegans cosmid ZC395.
A:Reference number: Z20385
A:Accession: T27538
A>Status: preliminary; translated from GB/EMBL/DBDUT
A:Molecule type: DNA
A:Residues: 1-433 <CON>
A:Cross-references: EMBL:U13642; PIDN:AB5386.1; GSPDB:GN00021; CESP:ZC395.6
A:Experimental source: strain Bristol N2, clone ZC395
C:Genetics:
A:Gene: CESP:ZC395.6
A:Map position: 3
A:Introns: 43/1; 201/1; 266/3; 310/2; 341/3; 377/3; 410/2

Alignment Scores:
Pred. No.: 4 88e-34 length: 433
Score: 531.50 Matches: 136
Percent Similarity: 50.34% Conservative: 84
Best Local Similarity: 31.12% Mismatches: 150
Query Match: 14.87% Indels: 67
DB: Gaps: 14

US-09-513-151-3 (1-2041) x T27538 (1-433)
OY 71 CGAACCCTACCTTTGATGATATCTCGGGGACCAGGGGACCGGCAATCCAGCTGGCG 130
Db 2 ArgthrAspproillellepheval1llegl1CysThrclYthrclYlsserAsprleugly 21
OY 131 TTGGAGCAGGCGCGGCGCTCGGGGAGATGCTGACGCTGACCTCATGACAGGCTAT 190
Db 22 ValalailealalyslsYrtyrlygllygluval1leserValaspsrsermetglnpheryr 41
OY 191 GAAGCCTAGACATCATCACCAACAGGTTTCTGCCCCAAGACAGAGAAATCTGCCGGCAC 250
Db 42 LysgileuasprrleatlnrAsnlysllethrclglnlglunleuserlyglnlyllegl1nls 61
OY 251 CACATGATCAGCTTTGGGATCCCTCT---GTAGCAATTAACACAGTGGTGGACTTCAGA 307
Db 62 HismetweterPheleuasnpProserGluuserfserYrthsnValh1sserPharg 81

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 18:21:20 ; Search time 47.8883 Seconds

(without alignments)
8194.498 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 3575
Sequence: 1 CTGCCATAGATGGCTCCG.....TTTACAGAAAAA 2041

Scoring table:

BLOSUM62	
Xgapop 10.0, Xgapext 0.5	
Ygapop 10.0, Ygapext 0.5	
Fgapop 6.0, Fgapext 7.0	
Delop 6.0, Delext 7.0	

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
-MODEL-frame+n2p model -DEV-rlh
-q=/cgn2_1/USPTO.spool/US09513151/unatc_15042003.141144_26390/app-query.fasta.1.2446
-DB=PIR-73 -QEXT-fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-Blosum62 -TRANS-human40.cdi -LIST=45
-DOCALLIN=200 -THR_SCORE-pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE-LOCAL
-OUTFMT=ptc -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09513151 @CIGN.1.1.58 @runat.15042003.141144.26390 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGIOCG -DEV-TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
1: PIR.73:*
2: PIR3:*
3: PIR3:*
4: PIR4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	548	15.3	428	2	667176
2	531.5	14.9	433	2	t27538
3	528.5	14.8	434	2	t38664
4	493	13.8	466	2	t52061
5	471	13.2	461	2	P84676
6	417	11.7	314	2	P83945
7	403.5	11.3	314	2	G69657
8	397.5	11.1	305	2	AF1236
9	395.5	11.1	305	2	AC1599
10	378.5	10.6	309	2	D97126
11	348	9.7	305	2	C72366
12	344.5	9.6	339	2	A71473
13	339	9.5	311	2	B89905
14	338	9.5	318	2	T05569

15	334.5	9.4	311	2	F64046
16	334	9.3	315	2	B82334
17	333	9.3	294	2	B86703
18	329.5	9.2	325	2	AB2464
19	328.5	9.2	323	2	E83028
20	327.5	9.2	314	2	C81735
21	324.5	9.1	311	2	S72942
22	324.5	9.1	314	2	F70505
23	324	9.1	311	2	D97945
24	322.5	9.0	357	2	F66708
25	321.5	9.0	313	2	E81879
26	320	9.0	294	2	G95077
27	319.5	8.9	342	2	E72019
28	319.5	8.9	342	2	DB6604
29	318.5	8.9	313	2	AD0046
30	314	8.8	305	2	G70391
31	313.5	8.8	330	2	T48100
32	312.5	8.7	317	2	H82848
33	311	8.7	316	2	B37318
34	310	8.7	316	2	C91272
35	310	8.7	316	2	C86113
36	305	8.5	312	2	T35111
37	302	8.4	316	2	AH1048
38	301.5	8.4	306	2	D70202
39	286.5	8.0	306	2	H75366
40	281.5	7.9	303	2	H97781
41	280	7.8	363	2	S75554
42	278.5	7.8	313	2	AB1142
43	275.5	7.7	326	2	G97604
44	272	7.6	298	2	B42643
45	272	7.6	298	2	A12826

ALIGNMENTS

RESULT 1

S67176
tRNA isopentenyltransferase (EC 2.5.1.8) - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein 05447w; protein YOR274w

C:Species: Saccharomyces cerevisiae

C:Date: 12-Jul-1996 #sequence revision 12-Jul-1996 #text change 21-Jul-2000

C:Accession: S67176; A26717; S72045

R:Cheret, G.; Sor, F.
submitted to the Protein Sequence Database, July 1996

A:Reference number: S67169

A:Accession: S67176

A:Molecule type: DNA

A:Residues: 1-428 <CHP>

A:Cross-references: EMBL:Z75182; NID:91420613; PID:e252418; PID:91420614; MIPS:YOR27

A:Experimental source: strain S288C

R:Naftarian, D.; Dihanich, M.E.; Martin, N.C.; Hopper, A.K.

Mol. Cell. Biol. 7, 185-191, 1987

A:Title: DNA sequence and transcript mapping of MOD5, features of the 5' region which

A:Reference number: A26717; MUID:87172703; PMID:3031457

A:Accession: A26717

A:Molecule type: DNA

A:Residues: 1-374, 'R', 376-428 <RNAJ>

A:Cross-references: EMBL:M15991

R:Cheret, G.; Bernardi, A.; Sor, F.

Yeast 12, 1059-1064, 1996

A:Title: DNA sequence analysis of the VPB1-SNF2 region on chromosome XV of Saccharom

A:Reference number: S72039; MUID:97051594; PMID:8896271

A:Accession: S72045

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-428 <CHP>

A:Cross-references: EMBL:X89633; NID:91279694; PID:CA61780.1; PID:91419759

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995

C:Genetics:

A:Gene: SGD:MOD5

A:Cross-references: MIPS:YOR274w; SGD:S0005800

A:Map position: 15R

C:Keywords: transferase

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 18:52:25 ; Search time 63.5 Seconds
(without alignments)

8565,812 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 661
Sequence: 1 CTGCCATAGATGGCTGCTCG.....TTTACAGAAAAA 2041

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1687582

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame+_n2p_model -DEV-xlh
-Q-/cgn2.1/USPTO.spool/US09513151/runat.15042003.141246.27205/app.query.fasta.1.2183
-DB-A-Geneseq.101002 -OFMT-fastan -SUFFIX-oln2p_rag -MINMATCH=0.1 -IOBCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-oligo -TRANS-human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=15 -MODE-LOCAL
-OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US09513151 -ECGN.1.1.72 -runat.15042003.141246.27205 -NCFU=6 -ICPU=3
-NO_XLPHY -NO_MMAR -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MAIN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A_Geneseq.101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	254	38.4	411	23	ABB97295
2	123	18.6	222	22	ABB10278
3	123	18.6	222	22	ABB10476
4	123	18.6	222	22	AAU23398
5	48	7.3	57	22	Novel human enzyme
6	48	7.3	57	22	Peptide #1530 enco
7	48	7.3	57	22	Peptide #1561 enco
8	48	7.3	57	22	Human brain expres
9	48	7.3	57	22	Human bone marrow
10	48	7.3	57	22	Peptide #1500 enco
11	16	2.4	221	20	Partial sequence o
12	9	1.4	94	22	Human colon cancer
13	8	1.2	11	19	Human tumour necro
14	8	1.2	11	19	Human tumour necro
15	8	1.2	11	19	Human tumour necro
16	8	1.2	12	19	Human tumour necro
17	8	1.2	48	22	Human immune/haema
18	8	1.2	60	22	Human digestive sy.
19	8	1.2	101	23	Human ORFX protein
20	8	1.2	118	22	Human polypeptide
21	8	1.2	120	22	Human quonoprotein
22	8	1.2	132	21	Mouse protein kina
23	8	1.2	136	23	Protonibacterium
24	8	1.2	212	22	Protonibacterium
25	8	1.2	216	20	Protonibacterium
26	8	1.2	219	22	Protonibacterium
27	8	1.2	234	22	Human secreted pro
28	8	1.2	269	22	Protonibacterium
29	8	1.2	277	22	Protonibacterium
30	8	1.2	330	23	Herbicidally activ
31	8	1.2	330	23	Herbicidally activ
32	8	1.2	351	22	Herbicidally activ
33	8	1.2	351	22	Herbicidally activ
34	8	1.2	357	23	Drosophila melanog
35	8	1.2	378	21	Herbicidally activ
36	8	1.2	378	21	Human lysophosphat
37	8	1.2	378	21	Membrane-bound pro
38	8	1.2	378	22	Human PRO1016 poly
39	8	1.2	378	22	Human ATP protein
40	8	1.2	378	22	Human lysophosphat
41	8	1.2	378	22	Human transferase
42	8	1.2	378	22	Human PRO1016 poly
43	8	1.2	378	22	Human PRO1016 poly
44	8	1.2	378	22	Human LIPAT delta
45	8	1.2	378	23	Human lysophosphat

ALIGNMENTS

RESULT 1
ABB97295
ID ABB97295 standard; Protein: 411 AA.

AC ABB97295;
XX 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 563.

XX Human; antiaiemic; vulnerary; antinflammatory; immunomodulator;
XX antinflammatory; cerebroprotective; cytostatic; gene therapy;
XX neuroprotective; antiparkinsonian; protein therapy; EST;
XX expressed sequence tag.

XX Homo sapiens.

XX OS
XX PN W0200222660-A2.

PD 21-MAR-2002.
XX 10-SEP-2001; 2001WO-US26015.
XX 11-SEP-2000; 2000US-0659671.
XX (HYSE-) HYSEQ INC.
PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
PI WPI; 2002-292408/33.
DR N-PSDB; ABN32481.
XX
PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.
XX
PS Example 2; SEQ ID NO 563; 509pp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC hematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention.
XX
SQ Sequence 411 AA;

Alignment Scores:
Pred. No.: 4,69e-242 Length: 411
Score: 254.00 Matches: 254
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.43% Indels: 0
DB: 23 Gaps: 0

US-09-513-151-3 (1-2041) x ABB97295 (1-411)
QY 179 ATGCGAGCTGTATGAAGGCTGACATCATCACACAAAGTTTCCCAAGACAGAGA 238
DB 1 MetGlnValIYrGlUGlYLeuAspIleleThraSnIyValSerAlaGlnGlnArg 20
QY 239 ATCTGCCGACACACATGATGACCTTGTGATCCTCTGTGACCAATTAACAGTGTG 298
DB 21 IleCysArgHISHisMetIleSerPheValAspProLeuValThraSnIyThraVal 40
QY 299 GACTTCAGAAATAGACAACTGCTGTGATTAAGATATATTGGCCGACAAATTCCT 358
DB 41 AspPheArgAsnArgAlaThraLeuIleGlnAspIlePheAlaArgAspIlePro 60
QY 359 ATTGTGGGGGAGGAACAATATATACATTAATCTGCTGCTGGAAGTTTCTCAT 418
DB 61 IleValValIGlYglYThraSnIyTyrIleGlnSerLeuIleTyrPlyValIleuValAsn 80
QY 419 ACCAAGCCCCAGAGATGGGCACTGAGAAAGTATTGACCGAAAGTGAGCTTGAAG 478
DB 81 ThrIysProGlnGlnIleMetGlyThrGlnIyValIleAspArgIyValGlnIleuGlnIy 100
QY 479 GAGGATGCTTGTACTTACAAACGCTTAAGCCAGGTGGACCCGAAATGGCTGCCAAG 538
DB 101 GluAspArgIyLeuValIleHisIyAspArgLeuSerGlnValAspProGlnMetAlaIalys 120
QY 539 CTGCACTCAACATGACAAAGCAAGTGGCCAGAGCTTGCAAGTTTGAAGAACAGAGA 598
DB 121 LeuHisProHisAspIyAspArgIyValAlaArgSerLeuGlnValPheGlnIleuArgIy 140
QY 599 ATCTGTACATGAAATTTCTCATGCTGCAACATACGGAAGAGGTGGTGGCTCCCTGGA 658
DB 141 IleSerHisSerGlnPheLeuHisArgGlnHisThrGlnGlnIyGlyGlyProLeuGlnIy 160

QY 659 GGTCCTGAGAGTTCCTTAACACCTTGACCTTGCTCATGCTGACACGACAGTTCTA 718
DB 161 GlyProLeuIySpHeserAsnProCysIleleuThraPheHisAlaAspGlnAlaValLeu 180
QY 719 GATGAGCGCTTGATTAAGAGGGTGATGACATGCTCTGCTGGCTTGTGAGAACTA 778
DB 181 AspIleuArgLeuAspIyAspValAspMetLeuAlaIalGlyLeuGlnIleu 200
QY 779 AGAATTTTCACACACCTTAATACAGAAATTTCCGGAATTAACCGAGACTATCAA 838
DB 201 ArgAspPheHisArgArgIyAsnGlnIyAsnValSerGlnAsnSerGlnAspIyGln 220
QY 839 CATGTAATCTTCACATCAATATGCGCTTCAGAGAAATTTACAGACTGATCACTAGAGCA 898
DB 221 HisIyIlePheGlnSerIleGlyPheIyGlnIleuHisGlnIyIleuIleThrGlnGly 240
QY 899 AAATGCACACTGAGAGACTAGTACCAAGCTTAAAGAAAGCA 940
DB 241 LysCysThrLeuGlnIleThraSerGlnIleuLeuIyIyGly 254

RESULT 2
ID ABB10278 standard; Protein: 222 AA.
XX ABB10278;
AC ABB10278;
XX 10-JAN-2002 (first entry)
XX
XX Human CDNA SEQ ID NO: 586.
DE
XX
XX Human; gene therapy; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW proliferative disorder; Inflammation.
XX Homo sapiens.
XX WO200154474-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01349.
XX
XX 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 24-FEB-2000; 2000US-184664P.
PR 02-MAR-2000; 2000US-186350P.
PR 16-MAR-2000; 2000US-189874P.
PR 17-MAR-2000; 2000US-190076P.
PR 18-APR-2000; 2000US-198123P.
PR 19-MAY-2000; 2000US-205155P.
PR 07-JUN-2000; 2000US-209467P.
PR 28-JUN-2000; 2000US-214886P.
PR 30-JUN-2000; 2000US-215135P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225213P.
PR 14-AUG-2000; 2000US-225214P.
PR 14-AUG-2000; 2000US-225266P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.

```
PR 14-AUG-2000; 2000US-225759P.
PR 18-AUG-2000; 2000US-226279P.
PR 22-AUG-2000; 2000US-226681P.
PR 22-AUG-2000; 2000US-226868P.
PR 22-AUG-2000; 2000US-227182P.
PR 23-AUG-2000; 2000US-227009P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 06-SEP-2000; 2000US-230437P.
PR 06-SEP-2000; 2000US-230438P.
PR 08-SEP-2000; 2000US-231242P.
PR 08-SEP-2000; 2000US-231243P.
PR 08-SEP-2000; 2000US-231413P.
PR 08-SEP-2000; 2000US-231414P.
PR 08-SEP-2000; 2000US-232080P.
PR 08-SEP-2000; 2000US-232081P.
PR 12-SEP-2000; 2000US-231968P.
PR 14-SEP-2000; 2000US-232397P.
PR 14-SEP-2000; 2000US-232398P.
PR 14-SEP-2000; 2000US-232399P.
PR 14-SEP-2000; 2000US-232400P.
PR 14-SEP-2000; 2000US-232401P.
PR 14-SEP-2000; 2000US-233063P.
PR 14-SEP-2000; 2000US-233064P.
PR 14-SEP-2000; 2000US-233065P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 25-SEP-2000; 2000US-234998P.
PR 26-SEP-2000; 2000US-235484P.
PR 27-SEP-2000; 2000US-235834P.
PR 27-SEP-2000; 2000US-235835P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239335P.
PR 13-OCT-2000; 2000US-239337P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241221P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241786P.
PR 20-OCT-2000; 2000US-241787P.
PR 20-OCT-2000; 2000US-241806P.
PR 20-OCT-2000; 2000US-241809P.
PR 20-OCT-2000; 2000US-241826P.
PR 01-NOV-2000; 2000US-244617P.
PR 08-NOV-2000; 2000US-246474P.
PR 08-NOV-2000; 2000US-246475P.
PR 08-NOV-2000; 2000US-246476P.
PR 08-NOV-2000; 2000US-246477P.
PR 08-NOV-2000; 2000US-246478P.
PR 08-NOV-2000; 2000US-246523P.
PR 08-NOV-2000; 2000US-246524P.
PR 08-NOV-2000; 2000US-246525P.
PR 08-NOV-2000; 2000US-246526P.
PR 08-NOV-2000; 2000US-246527P.
PR 08-NOV-2000; 2000US-246528P.
PR 08-NOV-2000; 2000US-246532P.
PR 08-NOV-2000; 2000US-246609P.
PR 08-NOV-2000; 2000US-246610P.

PR 08-NOV-2000; 2000US-246611P.
PR 08-NOV-2000; 2000US-246613P.
PR 17-NOV-2000; 2000US-249207P.
PR 17-NOV-2000; 2000US-249208P.
PR 17-NOV-2000; 2000US-249209P.
PR 17-NOV-2000; 2000US-249210P.
PR 17-NOV-2000; 2000US-249211P.
PR 17-NOV-2000; 2000US-249212P.
PR 17-NOV-2000; 2000US-249213P.
PR 17-NOV-2000; 2000US-249214P.
PR 17-NOV-2000; 2000US-249215P.
PR 17-NOV-2000; 2000US-249216P.
PR 17-NOV-2000; 2000US-249217P.
PR 17-NOV-2000; 2000US-249218P.
PR 17-NOV-2000; 2000US-249244P.
PR 17-NOV-2000; 2000US-249245P.
PR 17-NOV-2000; 2000US-249246P.
PR 17-NOV-2000; 2000US-249255P.
PR 17-NOV-2000; 2000US-249297P.
PR 17-NOV-2000; 2000US-249299P.
PR 17-NOV-2000; 2000US-249300P.
PR 01-DEC-2000; 2000US-250160P.
PR 01-DEC-2000; 2000US-250391P.
PR 05-DEC-2000; 2000US-251030P.
PR 05-DEC-2000; 2000US-251988P.
PR 05-DEC-2000; 2000US-256719P.
PR 06-DEC-2000; 2000US-251479P.
PR 06-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
PR 08-DEC-2000; 2000US-251989P.
PR 08-DEC-2000; 2000US-251990P.
PR 11-DEC-2000; 2000US-254097P.
PR 03-JAN-2001; 2001US-259678P.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-476161/51.
N-PSDB; ABA06500.

Isolated nucleic acid molecule encoding an inflammation-associated
polypeptide is used in preventing, treating or ameliorating a medical
condition
Claim 11; SEQ ID NO: 586; 859pp + Sequence Listing; English.
The present invention provides human cDNAs, proteins and related genomic
DNAs. These can be used in the treatment of neural, immune system,
muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders and inflammation. The present sequence
is a protein of the invention.
SQ Sequence 222 AA;

Alignment Scores:
Pred. No.: 1,88e-112 Length: 222
Score: 123.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.61% Indels: 0
DB: 22 Gaps: 0

US-09-513-151-3 (1-2041) x ABB10278 (1-222)
QY 2 TGCCATGAAGATGCGCTCCGTGCGACGAGTTCCTGTGGCAGTGGCCTCAGG 61
Db 1 CysHISLysMetAlaSerValAlaAlaAlaAlaAlaValAlaProValGlySerGlyLeuArg 20
QY 62 GGCCATGACAGGACCCCTACCTCTGTAGTATCTCGGGGCGACGCGGCAACGAAATCC 121
Db 21 GlyLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40
```

Qy	122	ACCGTGGGCTTGACAGCTGAGCGGAGCGGTGGAGATCGTCAGCGTCACTCATG	181
Db	41	ThiReuAlaLeuGlnIleuGlyGlnArgLeuGlyGlyIleValIleValSerAlaAspSerMet	60
Qy	182	CAGGTCTATGAAGCCCTAGACATCATCCACACAAAGCTTCTGCGCCAAAGAGAGATC	241
Db	61	GlnValIYrGIuGIuLeuAspIleIleThrAsnIysValSerAlaGlnGluGlnArgIle	80
Qy	242	TGCCGGCACCCATGATCAGCTTGTGTGATCTCTTGTGACAAATTACACAGTGTGGAC	301
Db	81	CysArgIhSIhSIhMetIleSerPheValAspProIeuValThrAsnYrThrValAsp	100
Qy	302	TTTCGAATATAGACAGACTGCTCTGATTTGAGAGATATATTGCGCGAGACAAATTCTAT	361
Db	101	PheArgAsnArgAlaIhAlhAlaLeuIleGluIhAspIlePheIhArgAspIlyIleProIle	120
Qy	362	GTGTGTGGA 370	
Db	121	ValValGly 123	
RESULT 3			
ABBI0476			
ID	ABBI0476	standard; Protein: 222 AA.	
AC	ABBI0476;		
XX			
DT	10-JAN-2002	(first entry)	
XX			
DE	Human cDNA SEQ ID NO: 784.		
XX			
KW	Human; gene therapy; neural disorder; immune system disorder;		
KM	muscular disorder; reproductive disorder; gastrointestinal disorder;		
KW	pulmonary disorder; cardiovascular disorder; renal disorder;		
XX	proliferative disorder; inflammation.		
OS	Homo sapiens.		
XX			
PN	MO200154474-A2.		
XX			
PD	02-AUG-2001.		
XX			
PF	17-JAN-2001; 2001MO-USO1349.		
XX			
PR	31-JAN-2000; 2000US-179065P.		
PR	04-FEB-2000; 2000US-180628P.		
PR	24-FEB-2000; 2000US-184664P.		
PR	02-MAR-2000; 2000US-186350P.		
PR	16-MAR-2000; 2000US-189874P.		
PR	17-MAR-2000; 2000US-190076P.		
PR	18-APR-2000; 2000US-198123P.		
PR	19-MAY-2000; 2000US-205515P.		
PR	07-JUN-2000; 2000US-209467P.		
PR	28-JUN-2000; 2000US-214886P.		
PR	30-JUN-2000; 2000US-215135P.		
PR	07-JUL-2000; 2000US-216647P.		
PR	07-JUL-2000; 2000US-216880P.		
PR	11-JUL-2000; 2000US-217487P.		
PR	11-JUL-2000; 2000US-217496P.		
PR	14-JUL-2000; 2000US-218290P.		
PR	26-JUL-2000; 2000US-220963P.		
PR	26-JUL-2000; 2000US-220964P.		
PR	14-AUG-2000; 2000US-224518P.		
PR	14-AUG-2000; 2000US-224519P.		
PR	14-AUG-2000; 2000US-224519P.		
PR	14-AUG-2000; 2000US-225213P.		
PR	14-AUG-2000; 2000US-225214P.		
PR	14-AUG-2000; 2000US-225266P.		
PR	14-AUG-2000; 2000US-225267P.		
PR	14-AUG-2000; 2000US-225268P.		
PR	14-AUG-2000; 2000US-225270P.		
PR	14-AUG-2000; 2000US-225447P.		
PR	14-AUG-2000; 2000US-225757P.		
PR	14-AUG-2000; 2000US-225758P.		
PR	14-AUG-2000; 2000US-225758P.		

PR	18-AUG-2000	2000US-225755P
PR	18-AUG-2000	2000US-226275P
PR	22-AUG-2000	2000US-226681P
PR	22-AUG-2000	2000US-226868P
PR	22-AUG-2000	2000US-227182P
PR	23-AUG-2000	2000US-227109P
PR	30-AUG-2000	2000US-228924P
PR	01-SEP-2000	2000US-228328P
PR	01-SEP-2000	2000US-229347P
PR	01-SEP-2000	2000US-229344P
PR	01-SEP-2000	2000US-229345P
PR	05-SEP-2000	2000US-229540P
PR	05-SEP-2000	2000US-229511P
PR	06-SEP-2000	2000US-230437P
PR	06-SEP-2000	2000US-230438P
PR	08-SEP-2000	2000US-231242P
PR	08-SEP-2000	2000US-231243P
PR	08-SEP-2000	2000US-231244P
PR	08-SEP-2000	2000US-231414P
PR	08-SEP-2000	2000US-231414P
PR	08-SEP-2000	2000US-233080P
PR	08-SEP-2000	2000US-233081P
PR	12-SEP-2000	2000US-233196P
PR	14-SEP-2000	2000US-233397P
PR	14-SEP-2000	2000US-233398P
PR	14-SEP-2000	2000US-233399P
PR	14-SEP-2000	2000US-233400P
PR	14-SEP-2000	2000US-233401P
PR	14-SEP-2000	2000US-233063P
PR	14-SEP-2000	2000US-233064P
PR	14-SEP-2000	2000US-233065P
PR	21-SEP-2000	2000US-234212P
PR	21-SEP-2000	2000US-234217P
PR	25-SEP-2000	2000US-234997P
PR	25-SEP-2000	2000US-234998P
PR	25-SEP-2000	2000US-235448P
PR	25-SEP-2000	2000US-235448P
PR	27-SEP-2000	2000US-235833P
PR	27-SEP-2000	2000US-235834P
PR	29-SEP-2000	2000US-236367P
PR	29-SEP-2000	2000US-236367P
PR	29-SEP-2000	2000US-236368P
PR	29-SEP-2000	2000US-236369P
PR	29-SEP-2000	2000US-236370P
PR	02-OCT-2000	2000US-236807P
PR	02-OCT-2000	2000US-237037P
PR	02-OCT-2000	2000US-237038P
PR	02-OCT-2000	2000US-237039P
PR	02-OCT-2000	2000US-237040P
PR	13-OCT-2000	2000US-239335P
PR	13-OCT-2000	2000US-239637P
PR	20-OCT-2000	2000US-240960P
PR	20-OCT-2000	2000US-241221P
PR	20-OCT-2000	2000US-241785P
PR	20-OCT-2000	2000US-241786P
PR	20-OCT-2000	2000US-241807P
PR	20-OCT-2000	2000US-241808P
PR	20-OCT-2000	2000US-241809P
PR	20-OCT-2000	2000US-241812P
PR	01-NOV-2000	2000US-244618P
PR	01-NOV-2000	2000US-244617P
PR	08-NOV-2000	2000US-246474P
PR	08-NOV-2000	2000US-246475P
PR	08-NOV-2000	2000US-246476P
PR	08-NOV-2000	2000US-246477P
PR	08-NOV-2000	2000US-246478P
PR	08-NOV-2000	2000US-246523P
PR	08-NOV-2000	2000US-246524P
PR	08-NOV-2000	2000US-246525P
PR	08-NOV-2000	2000US-246526P
PR	08-NOV-2000	2000US-246527P
PR	08-NOV-2000	2000US-246528P
PR	08-NOV-2000	2000US-246532P
PR	08-NOV-2000	2000US-246533P
PR	08-NOV-2000	2000US-246610P

QY	122	ACGTGGCGGTTGGACGACGAGCGCGGCTGGGGGTGAGATCGTCAGCGCTGACTCCATG	181
Db	41	ThrlenuAlaIeuGlnIeuGlyGlnArgIeuGlyGlnIleValSerAlaIaspSerMet	60
QY	182	CAGGTCAATTAAGGCCAGACATCATCCACCAAGAAGGTTTCTCCCAAGAGAGAAATC	241
Db	61	GlnValIYrIgluIgluLeuAspIleIleThrAsnIysValSerIleagIngIuIgluArgIle	80
QY	242	TGCGGACACACATGATCAGCTTTGTGGATCCTCTGTGACCAATTACACAGTGGTGAC	301
Db	81	CysArgHisHisMetIleSerPheValAspProIeuValThrAsnYrThrValValAsp	100
QY	302	TTTCGAATATGAGCAACTGCTGCGATTGAGATATATTTCGCCGAGACAAATTCCTATT	361
Db	101	PheArgAsnArgAlaIatrnAlaIeuIleGluAspIlePheAlaIargAspIleProIle	120
QY	362	GTGTGGGA 370	
Db	121	ValValGly 123	
RESULT 4			
AAU23398	ID	AAU23398 standard; Protein; 222 AA.	
AAU23398;	AC		
18-DEC-2001 (first entry)	DT		
Novel human enzyme polypeptide #484.	DE		
Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;	KW		
lyase; hyperproliferative disorder; immunodeficiency disorder;	KM		
autoimmune disorder; neurological disorder; metabolic disorder;	KW		
inflammatory disorder; cardiovascular disorder; reproductive disorder;	KW		
blood-related disorder; infectious disorder; cytostatic; anti arthritic;	KW		
nephrotropic; anticoagulant.	KW		
Homo sapiens.	OS		
MO200155301-A2.	XX		
02-AUG-2001.	XX		
17-JAN-2001; 2001WO-US01239.	PF		
31-JAN-2000; 2000US-0179065.	PR		
04-FEB-2000; 2000US-0180628.	PR		
24-FEB-2000; 2000US-0184664.	PR		
02-MAR-2000; 2000US-0186350.	PR		
16-MAR-2000; 2000US-0189874.	PR		
17-MAR-2000; 2000US-0190076.	PR		
18-APR-2000; 2000US-0198123.	PR		
19-MAY-2000; 2000US-0205515.	PR		
07-JUN-2000; 2000US-0209467.	PR		
28-JUN-2000; 2000US-0214886.	PR		
30-JUN-2000; 2000US-0215135.	PR		
07-JUL-2000; 2000US-0216647.	PR		
07-JUL-2000; 2000US-0216880.	PR		
11-JUL-2000; 2000US-0217487.	PR		
11-JUL-2000; 2000US-0217486.	PR		
14-JUL-2000; 2000US-0218290.	PR		
26-JUL-2000; 2000US-0220963.	PR		
26-JUL-2000; 2000US-0220964.	PR		
14-AUG-2000; 2000US-0224518.	PR		
14-AUG-2000; 2000US-0224519.	PR		
14-AUG-2000; 2000US-0225213.	PR		
14-AUG-2000; 2000US-0225214.	PR		
14-AUG-2000; 2000US-0225267.	PR		
14-AUG-2000; 2000US-0225266.	PR		
14-AUG-2000; 2000US-0225268.	PR		
14-AUG-2000; 2000US-0225270.	PR		
14-AUG-2000; 2000US-0225447.	PR		

PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233061.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0234984.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0244474.
 PR 08-NOV-2000; 2000US-0244475.
 PR 08-NOV-2000; 2000US-0244476.
 PR 08-NOV-2000; 2000US-0244477.
 PR 08-NOV-2000; 2000US-0244478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251865.
 PR 08-DEC-2000; 2000US-0251866.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCT INC.
 Rosen, CA, Barash SC, Ruben SM;
 WPI: 2001-465566/50.
 N-PSDB: AAS41268.
 Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases
 Claim 11: SEQ ID No 1394; 1180bp; English.
 The present invention relates to the isolation of novel human enzyme polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. influenza). The polynucleotides of the invention can also be used in gene therapy. AAS42915-AAS42814 represent the novel human enzyme polypeptides of the invention.
 Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 Sequence 222 AA:

Alignment Scores:

Pred. No.: 1.88e-112 Length: 222
 Score: 123.00 Matches: 123
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 18.61% Indels: 0
 DB: 22 Gaps: 0

US-09-513-151-3 (1-2041) x AAU23398 (1-222)

QY 2 TGCCATTAAGATGGCGTCCGCGGCTGCAGAGCAGTCTCTGTGGCAGTGGCTCAGG 61
 Db 1 CysHISLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
 QY 62 GGCCTGCACGAGGACCTTACCTCTTGTAGTATTCGCGGCGCAGCGGACCGCAATCC 121
 Db 21 GlyLeuGlnArgThrLeuProLeuValAlaIleLeuGlnAlaThrGlyThrGlyLysSer 40
 QY 122 ACGCTGGCGGTGACGCTAGGCGGCGGCTGCGGCTGAGATGTCAGGCTGACTCCATG 181
 Db 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGluIleValSerAlaLysPserMet 60
 QY 182 CAGGTCTATTAAGGCTTACATCATCATCAACAAGTTTGCCTGACCAAGCAGAGCAATC 241
 Db 61 GlnValTyrGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80
 QY 242 TGCCGCGCACCATGATCAGCTTTGTGATCCTCTTGTACCAATTACACAGTGTGAGC 301
 Db 81 CysArgHISHisMetIleSerPheValAspProLeuValThrAsnTyrThrValValAsp 100
 QY 302 TTCGAATATAGACACATGCTCTGTATTAAGATATATTGGCCGAGCAAAATTCCTATT 361
 Db 101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaIleAspLysIleProIle 120
 QY 362 GTGTGGGA 370
 Db 121 ValValGly 123

RESULT 5
 ABB28879
 ID ABB28879 standard; Peptide: 57 AA.
 XX
 AC ABB28879;
 DT 01-FEB-2002 (first entry)
 XX
 DE Peptide #1530 encoded by breast cell single exon nucleic acid probe.
 XX
 KW Human; microarray; single exon probe; gene expression; breast;
 KM disease; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157271-A2.
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001MO-US00662.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-46933/54.
 XX

PT New spatially-addressable set of single exon nucleic acid probes,
 PT useful for measuring gene expression in sample derived from human
 PT breast, comprises number of single exon nucleic acid probes
 XX
 XX Claim 27; SEQ ID NO 11847; 327bp + sequence listing; English.
 CC
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting
 CC the probes with a collection of detectably labelled nucleic acids
 CC derived from mRNA of human breast, and then measuring the label
 CC bound to each probe of the microarray. The probes are useful for
 CC verifying the expression of regions of genomic DNA predicted to
 CC encode proteins. They are useful for gene discovery, and for
 CC determining predisposition and/or prognosing breast disease. Gene
 CC expression analysis is useful for assessing the toxicity of chemical
 CC agents on cells. The microarray of this invention presents a far greater
 CC diversity of probes for measuring gene expression, with far less bias
 CC than expressed sequence tag microarrays. The method is suitable for
 CC rapid production of functional information from genomic sequence. The
 CC present sequence is a peptide encoded by a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
 CC
 XX
 S0 Sequence 57 AA:
 XX
 Alignment Scores:
 Pred. No.: 3.37e-38 Length: 57
 Score: 48.00 Matches: 48
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.26% Indels: 0
 DB: 22 Gaps: 0

US-09-513-151-3 (1-2041) x ABB28879 (1-57)
 QY 185 GTCTATGAAGGCTTACATCATCAACAAGTTTTCGCCAAGCAGAGATCTGC 244
 Db 1 ValTyrGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIleGys 20
 QY 245 CGGACACCATGATCAGCTTTGTGATCCTCTTGTGACCAATTACACAGTGTGACTTC 304
 Db 21 ArgHISHisMetIleSerPheValAspProLeuValThrAsnTyrThrValValAspPhe 40
 QY 305 AGAATAGACCACTGCTCTGATT 328
 Db 41 ArgAsnArgAlaThrAlaLeuIle 48

RESULT 6
 ABB34055
 ID ABB34055 standard; Peptide: 57 AA.
 XX
 AC ABB34055;
 DT 04-FEB-2002 (first entry)
 XX
 DE Peptide #1561 encoded by human foetal liver single exon probe.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.
 KM Homo sapiens.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001MO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR

PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human fetal liver -
 PS Claim 27; SEQ ID NO 26690; 659pp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 57 AA:

 Alignment Scores:
 Pred. No.: 3,37e-38 Length: 57
 Score: 48.00 Matches: 48
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.26% Indels: 0
 DB: 22 Gaps: 0

 US-09-513-151-3 (1-2041) x ABB34055 (1-57)

 QY 185 GTCRTATGAGGCTTGAGATCATCCACCAAGGTTTCTGCCAAGACAGAGATCTGC 244
 DB 1 ValtyrGluglyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIleCys 20
 QY 245 CGGCACCATGATGATGAGTCTTGATCCTCTGTGACCAATTACACAGTGAGACTTC 304
 DB 21 ArgHSHSmEtIleSerPheValAspProLeuValThrAsnTyThValValAspPhe 40
 QY 305 AGAATAGAGCAACTGCTCTGATT 328
 DB 41 ArgAsnArgAlaThrAlaLeuIle 48

 RESULT 7
 AAM54837
 ID AAM54837 standard; Protein: 57 AA.
 XX
 AC AAM54837;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 26942.
 XX
 KW Human: brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 XX brains -
 PS Example 4; SEQ ID NO: 26942; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 XX
 SQ Sequence 57 AA:

 Alignment Scores:
 Pred. No.: 3,37e-38 Length: 57
 Score: 48.00 Matches: 48
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.26% Indels: 0
 DB: 22 Gaps: 0

 US-09-513-151-3 (1-2041) x AAM54837 (1-57)

 QY 185 GTCRTATGAGGCTTGAGATCATCCACCAAGGTTTCTGCCAAGACAGAGATCTGC 244
 DB 1 ValtyrGluglyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIleCys 20
 QY 245 CGGCACCATGATGATGAGTCTTGATCCTCTGTGACCAATTACACAGTGAGACTTC 304
 DB 21 ArgHSHSmEtIleSerPheValAspProLeuValThrAsnTyThValValAspPhe 40
 QY 305 AGAATAGAGCAACTGCTCTGATT 328
 DB 41 ArgAsnArgAlaThrAlaLeuIle 48

 RESULT 8
 AAM67220
 ID AAM67220 standard; Protein: 57 AA.
 XX
 AC AAM67220;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 27526.
 XX
 KW Human: bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 XX
 OS Homo sapiens.
 XX
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00668.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 27526; 658bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 CC
 XX
 SQ Sequence 57 AA:
 Alignment Scores:
 Pred. No.: 3.37e-38 Length: 57
 Score: 48.00 Matches: 48
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 7.26% Gaps: 0
 DB: 22
 US-09-513-151-3 (1-2041) x AAM67220 (1-57)
 QY 185 GTCATGAGGCTAGACATCATCCACCAAGGTTTCTGCCAGACAGAGAAATCTGC 244
 Db 1 ValTyrgluGlyLeuAspIleIleThrasnlySValSerAlaGluGlnArgIleCys 20
 QY 245 CGGCACACATGATCAGCTTTGTGATCCTCTTGACCAATTAACAGTGTGACCTTC 304
 Db 21 ArgHSHSMetIleSerPheValAspProLeuValThrAsnTyThrValValAspPhe 40
 QY 305 AGAAATAGAGCAACTGCTCTGATT 328
 Db 41 ArgAsnArgAlaThrAlaLeuIle 48
 RESULT 9
 AAM15066
 ID AAM15066 standard; Protein: 57 AA.
 AC AAM15066;
 XX
 DT 12-OCT-2001 (first entry)
 DE Peptide #1500 encoded by probe for measuring cervical gene expression.
 XX
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157278-A2.
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00670.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells -
 XX
 PS Claim 27; SEQ ID No 19692; 487bp; English.
 XX
 CC The present invention relates to human single exon nucleic acid probes
 CC (SENP; see A110068-A128459). The present sequence is a peptide encoded
 CC by one such probe. The SENPs are derived from human Hela cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer.
 CC
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 57 AA:
 Alignment Scores:
 Pred. No.: 3.37e-38 Length: 57
 Score: 48.00 Matches: 48
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 7.26% Gaps: 0
 DB: 22
 US-09-513-151-3 (1-2041) x AAM15066 (1-57)
 QY 185 GTCATGAGGCTAGACATCATCCACCAAGGTTTCTGCCAGACAGAGAAATCTGC 244
 Db 1 ValTyrgluGlyLeuAspIleIleThrasnlySValSerAlaGluGlnArgIleCys 20
 QY 245 CGGCACACATGATCAGCTTTGTGATCCTCTTGACCAATTAACAGTGTGACCTTC 304
 Db 21 ArgHSHSMetIleSerPheValAspProLeuValThrAsnTyThrValValAspPhe 40
 QY 305 AGAAATAGAGCAACTGCTCTGATT 328
 Db 41 ArgAsnArgAlaThrAlaLeuIle 48
 RESULT 10
 AAM27510
 ID AAM27510 standard; Protein: 57 AA.
 AC AAM27510;
 XX
 DT 17-OCT-2001 (first entry)
 DE Peptide #1547 encoded by probe for measuring placental gene expression.
 XX
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200157272-A2.
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00663.
 XX
 PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
PS
XX Claim 27; SEQ ID No 27779; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENPs;
CC see AAT31315-AAT57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 57 AA:

Alignment Scores:
Pred. No.: 3,37e-38 Length: 57
Score: 48.00 Matches: 48
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7,268 Indels: 0
DB: 22 Gaps: 0

US-09-513-151-3 (1-2041) x AAM27510 (1-57)

OY 185 GTCTATGAAGGCTGACATCATCAACAAGGTTTCTGCCCAAGAGAGAGATCTGC 244
DB 1 VAIYTGUGLYLenuaspillelThrasnlysvaSerAlaGlnGlnArgrilecys 20
OY 245 CGGCACACATGATCAGCTTGTGGATCCTCTGTGACCAATTACAGAGTGGAGCTTC 304
DB 21 ArgHSHSMetIleSerPheValaspProleuValThrAsnTyThValValasphe 40
OY 305 AGAATAGAGCAACTGCTCTGATT 328
DB 41 ArgAsnArgAlaThrAlaLeulle 48

RESULT 11

ID AAY02534 standard; Protein: 221 AA.

XX AAY02534;

XX 15-JUL-1999 (first entry)

DE Partial sequence of the Drosophila homologue of gro-1.

KW gro-1 operon; gro-1 gene; gop-1 gene; gop-2 gene; gop-2 gene;
KW hab-1 gene; cancer; aging; longevity; tumour formation;
KW physiological clock.

XX Drosophila sp.

XX WO9910482-A1.

XX 04-MAR-1999.

XX 20-AUG-1998; 98WO-CA00803.

XX 25-AUG-1997; 97CA-2210251.

XX

PA (UYMC-) UNIV MCGILL.

XX Barnes T, Hekimi S, Lakowski B, Lemieux J;

XX WPI; 1999-190615/16.

PT Molecular identity of the gro-1 gene - useful for cancer diagnosis
PT and/or prognosis, and where compounds affecting encoded proteins are
PT useful for enhancing longevity of a host and inhibiting tumour
PT formation

XX Disclosure; Fig 10; 93pp; English.

XX
XX The specification describes the five genes of the *Caenorhabditis elegans*
CC gro-1 operon (AA336071). The operon contains the gro-1 gene (AA336072),
CC the gop-1 gene (AA336074), the gop-2 gene (AA336075), the gop-3 gene
CC (AA336076), and the hap-1 gene (AA336077). The gro-1 gene can be used in
CC a method for the diagnosis and/or prognosis of cancer in a patient.
CC Transgenic mice containing a gene knock-out of a murine gene homologue
CC of the gro-1 gene are useful as models of aging and cancer. The
CC proteins encoded by the genes are useful for identifying compounds
CC that affect the enzymatic activity of these proteins, in order to
CC enhance longevity of a host and inhibit tumour formation. The gro-1
CC gene, together with the gop-1, gop-2, gop-3 and hap-1 genes enables
CC study of a physiological clock. The present sequence represents a
CC partial sequence of the *Drosophila* homologue of gro-1.

SQ Sequence 221 AA:

Alignment Scores:
Pred. No.: 1,28e-06 Length: 221
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2,42% Indels: 0
DB: 20 Gaps: 0

US-09-513-151-3 (1-2041) x AAY02534 (1-221)

OY 356 CCTATTGTTGGAGAGAACCAATTATTGATGATCTGCTCTGG 403
DB 137 ProIleValIaIglYThrasnTyrrIleGlnSerLeuLeuTrp 152

RESULT 12

AA674967

ID AAG74967 standard; Protein: 94 AA.

XX AAG74967;

XX 03-SEP-2001 (first entry)

DE Human colon cancer antigen protein SEQ ID NO:5731.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

XX colorectal carcinoma.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.

XX 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI; 2001-235357/24.
XX N-PSDB; AAH34372.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
PT
XX
PS Claim 11; Page 7257; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 94 AA:
Alignment Scores:
Pred. No.: 11.9 Length: 94
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 22 Gaps: 0
US-09-513-151-3 (1-2041) x AAG74967 (1-94)
QY 1017 CTGCTCTTGAATCGTCAAGTTTCA 1043
DB 11 LeuLeuLeuLysSerCysValSer 19
RESULT 13
AAW47911
ID AAW47911 standard; peptide: 11 AA.
XX
AC AAW47911;
XX
DT 09-JUN-1998 (first entry)
XX
DE Human tumour necrosis factor derived peptide 1168.
XX
KW Human: tumour necrosis factor; neutrophil; monocyte; macrophage;
KW stimulatory activity; immunosuppressant; cancer; infection.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9748725-A1.
XX
PD 24-DEC-1997.
XX
PF 20-JUN-1997; 97WO-AU00395.
XX
PR 29-OCT-1996; 96AU-0003309.
PR 21-JUN-1996; 96AU-0000610.
PR 06-SEP-1996; 96AU-0002165.
XX
XX (PEPT-) PEPTTECH LTD.
XX
PI Mack PC, Rathjen DA, Sleight JM, Widmer F.
XX
XX WPI; 1998-063077/06.

XX Tumour necrosis factor derived peptide(s) having neutrophil and/or
PT monocyte-macrophage stimulatory activity - used for treating e.g.
PT infections, immunosuppression or cancers
XX
PS Claim 4; Page 44; 65pp; English.
XX
CC The present sequence represents a peptide derived from a human tumour
CC necrosis factor peptide. The peptide can have neutrophil and/or
CC monocyte/macrophage stimulatory activity. The peptide is derived from
CC the sequence of tumour necrosis factor (TNF)-derived peptide 419
CC (PETHYLRHRI; see AU74762/91 and 44664/93). It can have improved
CC properties such as increased potency, extended in vivo half life or,
CC particularly, specificity of action. It can have neutrophil stimulatory
CC activity (class 1), equal neutrophil and monocyte/macrophage stimulatory
CC activity (class 2), or preferentially enhanced monocyte/macrophage
CC stimulatory activity (class 3). It can be used for the treatment or
CC prevention of infections. In particular, it can be used for treating
CC AIDS, cancer, diabetes, nosocomial infection, tuberculosis, cystic
CC fibrosis, community acquired pneumonia, meningitis, Mycobacteria,
CC Chlamydia, Brucellae, Francisella, Pasteurella, Legionellosis,
CC Histoplasmosis, Listeriosis, Pneumocystis carinii, Trypanosoma cruzi,
CC coccidian parasitical infection, an inherited primary neutropenic
CC disorder, an inherited defect of phagocytic cell function, an acquired
CC defect of phagocytic cell function, immunosuppression due to the
CC administration of immunosuppressive drugs, and other bacterial, fungal,
CC viral or protozoan infection, infectious mononucleosis, paroxysmal
CC nocturnal, haemoglobinuria, leukaemia, lymphoma, myelofibrosis, or graft
CC versus host disease.
XX
SQ Sequence 11 AA:
Alignment Scores:
Pred. No.: 150 Length: 11
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.22% Indels: 0
DB: 19 Gaps: 0
US-09-513-151-3 (1-2041) x AAW47911 (1-11)
QY 1967 ACAGCAGTGATATACACACACA 1944
DB 3 ThrAlaValIleIleThrHisPThr 10
RESULT 14
AAW47917
ID AAW47917 standard; peptide: 11 AA.
XX
AC AAW47917;
XX
DT 09-JUN-1998 (first entry)
XX
DE Human tumour necrosis factor derived peptide 1175.
XX
KW Human: tumour necrosis factor; neutrophil; monocyte; macrophage;
KW stimulatory activity; immunosuppressant; cancer; infection.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9748725-A1.
XX
PD 24-DEC-1997.
XX
PF 20-JUN-1997; 97WO-AU00395.
XX
XX Key Location/Qualifiers
XX Modified-site 1
XX /note= "nipoctic acid"

```

PR 29-OCT-1996; 96AU-0003309.
PR 21-JUN-1996; 96AU-0000610.
PR 06-SEP-1996; 96AU-0002165.
XX
XX (PEPT-) PEPTech LTD.
XX
PI Mack PO, Rathjen DA, Sleigh JM, Widmer F;
XX
XX WPI: 1998-063077/06.
XX
XX Tumour necrosis factor derived peptide(s) having neutrophil and/or
XX monocyte-macrophage stimulatory activity - used for treating e.g.
XX infections, immunosuppression or cancers
XX
XX Claim 4; Page 44; 65pp; English.
XX
XX The present sequence represents a peptide derived from a human tumour
XX necrosis factor peptide. The peptide can have neutrophil and/or
XX monocyte/macrophage stimulatory activity. The peptide is derived from
XX the sequence of tumour necrosis factor (TNF)-derived peptide 419
XX (PSTHVLITHTI; see AU74762/91 and 44664/93). It can have improved
XX properties such as increased potency, extended in vivo half life or,
XX particularly, specificity of action. It can have neutrophil stimulatory
XX activity (class 1), equal neutrophil and monocyte/macrophage stimulatory
XX activity (class 2), or preferentially enhanced monocyte/macrophage
XX stimulatory activity (class 3). It can be used for the treatment or
XX prevention of infections. In particular, it can be used for treating
XX AIDS, cancer, diabetes, nosocomial infection, tuberculosis, cystic
XX fibrosis, community acquired pneumonia, meningitis, Mycobacteria,
XX Chlamydia, Brucellae, Francisella, Pasteurellosis, Legionellosis,
XX Histoplasmosis, listeriosis, Pneumocystis carni, Trypanosoma cruzi,
XX coccidian parasitical infection, an inherited primary neutropenic
XX disorder, an inherited primary defect of phagocytic cell function, an
XX inherited secondary defect of phagocytic cell function, an acquired
XX defect of phagocytic cell function, immunosuppression due to the
XX administration of immunosuppressive drugs, and other bacterial, fungal,
XX viral or protozoan infection, infectious mononucleosis, paroxysmal
XX nocturnal, haemoglobinuria, leukaemia, lymphoma, myelofibrosis, or graft
XX versus host disease.
XX
XX Sequence 11 AA:
XX
XX Alignment Scores:
XX
XX Pred. No.: 150 Length: 11
XX Score: 8.00 Matches: 8
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 1.22% Indels: 0
XX DB: Gaps: 0
XX
XX US-09-513-151-3 (1-2041) x AAM47917 (1-11)
XX
XX QY 1967 ACAGCAGTGATATACTACACACA 1944
XX |||||||||||||||||||
XX DB 3 ThrAlaValIleIleThrHisThr 10
XX
XX RESULT 15
XX AAM47901
XX ID AAM47901 standard; peptide; 11 AA.
XX
XX AC AAM47901;
XX
XX XX
XX DT 09-JUN-1998 (first entry)
XX
XX DE Human tumour necrosis factor derived peptide 1100.
XX
XX XX
XX KW Human; tumour necrosis factor; neutrophil; monocyte; macrophage;
XX stimulatory activity; immunosuppressant; cancer; infection.
XX
XX OS Synthetic.
XX OS Homo sapiens.
XX
XX PN WO9748725-A1.

```

```

XX
XX PD 24-DEC-1997.
XX
XX XX
XX FE 20-JUN-1997; 97WO-AU00395.
XX
XX PR 29-OCT-1996; 96AU-0003309.
XX PR 21-JUN-1996; 96AU-0000610.
XX PR 06-SEP-1996; 96AU-0002165.
XX
XX (PEPT-) PEPTech LTD.
XX
XX Mack PO, Rathjen DA, Sleigh JM, Widmer F;
XX
XX WPI: 1998-063077/06.
XX
XX Tumour necrosis factor derived peptide(s) having neutrophil and/or
XX monocyte-macrophage stimulatory activity - used for treating e.g.
XX infections, immunosuppression or cancers
XX
XX Claim 4; Page 44; 65pp; English.
XX
XX The present sequence represents a peptide derived from a human tumour
XX necrosis factor peptide. The peptide can have neutrophil and/or
XX monocyte/macrophage stimulatory activity. The peptide is derived from
XX the sequence of tumour necrosis factor (TNF)-derived peptide 419
XX (PSTHVLITHTI; see AU74762/91 and 44664/93). It can have improved
XX properties such as increased potency, extended in vivo half life or,
XX particularly, specificity of action. It can have neutrophil stimulatory
XX activity (class 1), equal neutrophil and monocyte/macrophage stimulatory
XX activity (class 2), or preferentially enhanced monocyte/macrophage
XX stimulatory activity (class 3). It can be used for the treatment or
XX prevention of infections. In particular, it can be used for treating
XX AIDS, cancer, diabetes, nosocomial infection, tuberculosis, cystic
XX fibrosis, community acquired pneumonia, meningitis, Mycobacteria,
XX Chlamydia, Brucellae, Francisella, Pasteurellosis, Legionellosis,
XX Histoplasmosis, listeriosis, Pneumocystis carni, Trypanosoma cruzi,
XX coccidian parasitical infection, an inherited primary neutropenic
XX disorder, an inherited primary defect of phagocytic cell function, an
XX inherited secondary defect of phagocytic cell function, an acquired
XX defect of phagocytic cell function, immunosuppression due to the
XX administration of immunosuppressive drugs, and other bacterial, fungal,
XX viral or protozoan infection, infectious mononucleosis, paroxysmal
XX nocturnal, haemoglobinuria, leukaemia, lymphoma, myelofibrosis, or graft
XX versus host disease.
XX
XX Sequence 11 AA:
XX
XX Alignment Scores:
XX
XX Pred. No.: 150 Length: 11
XX Score: 8.00 Matches: 8
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 1.22% Indels: 0
XX DB: Gaps: 0
XX
XX US-09-513-151-3 (1-2041) x AAM47901 (1-11)
XX
XX QY 1967 ACAGCAGTGATATACTACACACA 1944
XX |||||||||||||||||||
XX DB 3 ThrAlaValIleIleThrHisThr 10
XX
XX Search completed: April 21, 2003, 19:14:30
XX Job time : 66.5 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 21, 2003, 18:51:16 ; Search time 13.7893 Seconds

(without alignments)
8416.087 Million cell updates/sec

Title: US-09-513-151-3_COPY_1121_1210

Perfect score: 164

Sequence: 1 TGTGACCTCTGTGTCGAAAT.....TGAACCAACTGAGAGAAAAGA 90

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
	Ygapop 10.0 , Ygapext 0.5
	Fgapop 6.0 , Fgapext 7.0
	Delop 6.0 , Delext 7.0

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 9138288

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

```

-MODEL=frame_plus.n2p.model -DEV=xlh
-O/cgnt2_1/USPRO.spool/US09513151.runtat_15042003_141145_26426/app_query.fasta_1.2446
-DB=patents_AA_Main.qfmt.fasta -SUFFIX=n2p.ram -MIMATCH=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62
-TRANS=numa4.cdd -LIST=45 -DOCALIGN=200 -THR_SCORE=pcr -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09513151 @CGN1.1.422 @runtat_15042003_141145_26426
-NCPU=6 -ICPU=3 -NO MAP -LARGESQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -HEADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database :

Pending_Patents_AA_Main:

1:	/cgnt2_6/prodata/1/paa/PCRTUS.COMB.pep.*
2:	/cgnt2_6/prodata/1/paa/US06.COMB.pep.*
3:	/cgnt2_6/prodata/1/paa/US07.COMB.pep.*
4:	/cgnt2_6/prodata/1/paa/US08.COMB.pep.*
5:	/cgnt2_6/prodata/1/paa/US081.COMB.pep.*
6:	/cgnt2_6/prodata/1/paa/US082.COMB.pep.*
7:	/cgnt2_6/prodata/1/paa/US083.COMB.pep.*
8:	/cgnt2_6/prodata/1/paa/US084.COMB.pep.*
9:	/cgnt2_6/prodata/1/paa/US085.COMB.pep.*
10:	/cgnt2_6/prodata/1/paa/US086.COMB.pep.*
11:	/cgnt2_6/prodata/1/paa/US087.COMB.pep.*
12:	/cgnt2_6/prodata/1/paa/US088.COMB.pep.*
13:	/cgnt2_6/prodata/1/paa/US089.COMB.pep.*
14:	/cgnt2_6/prodata/1/paa/US090.COMB.pep.*
15:	/cgnt2_6/prodata/1/paa/US091.COMB.pep.*
16:	/cgnt2_6/prodata/1/paa/US092.COMB.pep.*
17:	/cgnt2_6/prodata/1/paa/US093.COMB.pep.*
18:	/cgnt2_6/prodata/1/paa/US094.COMB.pep.*
19:	/cgnt2_6/prodata/1/paa/US095.COMB.pep.*
20:	/cgnt2_6/prodata/1/paa/US096.COMB.pep.*
21:	/cgnt2_6/prodata/1/paa/US097.COMB.pep.*
22:	/cgnt2_6/prodata/1/paa/US098.COMB.pep.*
23:	/cgnt2_6/prodata/1/paa/US099.COMB.pep.*
24:	/cgnt2_6/prodata/1/paa/US100.COMB.pep.*
25:	/cgnt2_6/prodata/1/paa/US101.COMB.pep.*
26:	/cgnt2_6/prodata/1/paa/US102.COMB.pep.*
27:	/cgnt2_6/prodata/1/paa/US60.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	164	100.0	143	21	US-09-757-028-2199
2	164	100.0	143	26	US-10-222-911-2199
3	164	100.0	465	1	PCT-US02-07826-151
4	164	100.0	465	24	US-10-097-940-151
5	164	100.0	467	1	PCT-US02-07826-153
6	164	100.0	467	24	US-10-097-940-153
7	64	39.0	294	27	US-60-389-987-773
8	64	39.0	294	27	US-60-412-418-773
9	64	39.0	635	27	US-60-167-217-10879
10	64	39.0	884	27	US-60-173-464-8824
11	63	38.4	617	1	PCT-US01-18569-3108
12	63	38.4	617	26	US-10-264-049-3108
13	63	38.4	845	26	US-10-205-219-133
14	63	38.4	847	23	US-09-976-594-401
15	63	38.4	847	25	US-10-153-668-450
16	63	38.4	847	27	US-60-240-409-401
17	63	38.4	864	1	PCT-US01-08631-49060
18	61.5	37.5	354	25	US-10-179-131-10074
19	61.5	37.5	386	16	US-09-248-796-17815
20	61.5	37.5	386	27	US-60-096-409-17815
21	61	37.2	314	20	US-09-614-150-24477
22	61	37.2	314	27	US-60-191-637-24572
23	61	37.2	314	27	US-60-191-681-19349
24	60	36.6	251	26	US-10-219-999-57395
25	60	36.6	289	21	US-09-708-427-56054
26	60	36.6	324	21	US-09-708-427-56053
27	60	36.6	430	19	US-09-513-151-2
28	60	36.6	466	21	US-09-708-427-56052
29	59	36.0	428	21	US-09-513-956-73087
30	58	35.8	72	19	US-09-513-956-73087
31	58	35.8	72	19	US-09-513-956-73087
32	58	35.8	121	19	US-09-513-956-73086
33	58	35.8	121	19	US-09-513-956-73086
34	58	35.4	578	20	US-09-614-150-1971
35	58	35.4	578	27	US-60-191-637-1977
36	58	35.4	578	27	US-60-191-681-1578
37	58	35.4	884	20	US-09-614-150-10866
38	58	35.4	884	27	US-60-191-637-10898
39	58	35.4	884	27	US-60-191-681-8546
40	57.5	35.1	935	20	US-09-614-150-1641
41	57.5	35.1	935	27	US-60-173-464-22515
42	57.5	35.1	935	27	US-60-191-637-1647
43	57.5	35.1	935	27	US-60-191-681-1323
44	57.5	35.1	956	27	US-60-161-932-1331
45	57	35.2	126	21	US-09-758-472-4966

ALIGNMENTS

RESULT 1
US-09-757-028-2199
Sequence 2199, Application US/09757028
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PM001
CURRENT FILING DATE: 2001-01-09
CURRENT APPLICATION NUMBER: US/09/757,028
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/180,628
NUMBER OF SEQ ID NOS: 2660
SOFTWARE: PatentIn Ver. 2.0

```

: SEQ ID NO 2199
: LENGTH: 143
: TYPE: PROT
: ORGANISM: Homo sapiens
US-09-757-028-2199

Alignment Scores:
Pred. No.: 1.97e-15 Length: 143
Score: 164.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-09-513-151-3_COPY.1121_1210 (1-90) x US-09-757-028-2199 (1-143)
QY 1 TGTGACCTGTGTGATCATCATCATCTGTTGGGATCGCGAATGGGCGACGACATAAATCC 60
|||||
Db 73 CysaapleucysaSPARIIlellelleelgysapra:gIutrpAlaIahisIleYsser 92
|||||

QY 61 AATCCCACTTGAAACCAACTGAAGAAAGA 90
|||||
Db 93 LysSerHisLeuAsnGlnLeuIuIysaY 102
|||||

RESULT 3
: Sequence 151, Application PC/TUS0207826
: GENERAL INFORMATION:
: APPLICANT: Millennium Pharmaceuticals, Inc. et al.
: TITLE OF INVENTION: Nucleic Acid Molecules and Proteins for The Identification
: FILE REFERENCE: MRI-030PC

: SEQ ID NO 2199
: LENGTH: 143
: TYPE: PROT
: ORGANISM: Homo sapiens
US-10-222-911-2199

Alignment Scores:
Pred. No.: 1.97e-15 Length: 143
Score: 164.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 26 Gaps: 0

US-09-513-151-3_COPY.1121_1210 (1-90) x US-10-222-911-2199 (1-143)
QY 1 TGTGACCTGTGTGATCATCATCATCTGTTGGGATCGCGAATGGGCGACGACATAAATCC 60
|||||
Db 73 CysaapleucysaSPARIIlellelleelgysapra:gIutrpAlaIahisIleYsser 92
|||||

QY 61 AATCCCACTTGAAACCAACTGAAGAAAGA 90
|||||
Db 93 LysSerHisLeuAsnGlnLeuIuIysaY 102
|||||

RESULT 2
: Sequence 2199, Application US/10222911
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PM001CIN
: CURRENT APPLICATION NUMBER: US/10/222,911
: PRIOR FILING DATE: 2002-08-19
: PRIOR APPLICATION NUMBER: 09/757,028
: PRIOR FILING DATE: 2001-01-09
: PRIOR APPLICATION NUMBER: 60/179,065
: PRIOR FILING DATE: 2000-01-31
: PRIOR APPLICATION NUMBER: 60/180,628
: PRIOR FILING DATE: 2000-02-04
: NUMBER OF SEQ ID NOS: 2660
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2199
: LENGTH: 143
: TYPE: PROT
: ORGANISM: Homo sapiens
US-10-222-911-2199

```

```

CURRENT APPLICATION NUMBER: PCT-US02/07826
CURRENT FILING DATE: 2002-03-14
PRIORITY APPLICATION NUMBER: 60/276,025
PRIORITY FILING DATE: 2001-03-14
PRIORITY APPLICATION NUMBER: 60/325,149
PRIORITY FILING DATE: 2001-09-27
PRIORITY APPLICATION NUMBER: 60/276,026
PRIORITY FILING DATE: 2001-03-14
PRIORITY APPLICATION NUMBER: 60/324,967
PRIORITY FILING DATE: 2001/09/26
PRIORITY APPLICATION NUMBER: 60/311,732
PRIORITY FILING DATE: 2001-08-10
PRIORITY APPLICATION NUMBER: 60/325,102
PRIORITY FILING DATE: 2001-09-26
PRIORITY APPLICATION NUMBER: 60/323,580
PRIORITY FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 151
LENGTH: 465
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-07826-151

Alignment Scores:
Pred. No.: 2,17e-15 Length: 465
Score: 164.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x PCT-US02-07826-151 (1-465)
Oy 1 TGTGACCTGTGCATCATCATTCATTGGGGAGTCCGAATGGCGACCACTAAATCC 60
Db 335 CysaspleCysasphrgllellelleglyaspargglutPhlaiahislieysSer 414
Oy 61 AAATCCACTTGACCACCACTGAAGAAGA 90
Db 415 LysserHisLeuAsnGlnIleuLysIAsTy 424

RESULT 4
US-10-097-340-151
Sequence 151, Application US/10097340
GENERAL INFORMATION:
APPLICANT: John MONAHAN
APPLICANT: Manjula GANNAVARAPU
APPLICANT: Sebastian HOERSCH
APPLICANT: Shubhangi KAMATKAR
APPLICANT: Steve G. KOVATS
APPLICANT: Rachel E. MEYERS
APPLICANT: Michael MORRISSEY
APPLICANT: Peter OLANDT
APPLICANT: Aml SEN
APPLICANT: Peter VEIBY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. BAST, Jr.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHMANDT
APPLICANT: Xumei ZHAO
APPLICANT: Karen GLAYT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification
FILE REFERENCE: MRI-030
FILE REFERENCE: Assessment, Prevention, and Therapy of Ovarian Cancer
CURRENT APPLICATION NUMBER: US/10/097,340
CURRENT FILING DATE: 2002-03-14
PRIORITY APPLICATION NUMBER: 60/276,025
PRIORITY FILING DATE: 2001-03-14
PRIORITY APPLICATION NUMBER: 60/325,149
PRIORITY FILING DATE: 2001-09-26
PRIORITY APPLICATION NUMBER: 60/276,026
PRIORITY FILING DATE: 2001-03-14

```

```

; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-151

Alignment Scores:
Pred. No.: 2,17e-15      Length: 465
Score: 164.00           Matches: 30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 24                      Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-097-340-151 (1-465)
OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGGAGCGACGACATAAAATCC 60
Db 395 Cysaspleucysaspargillellelleglyaspargglutrpalaahisileysse 414
OY 61 AAATCCCACTTGAAACCACTGAAGAAAGA 90
Db 415 LysSerHisLeuansGlnLeuLysLysArg 424

RESULT 5
PCT-US02-07826-153
; Sequence 153, Application PC/TUS0207826
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030PC
; CURRENT APPLICATION NUMBER: PCT/US02/07826
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-07826-153

Alignment Scores:
Pred. No.: 2,17e-15      Length: 467
Score: 164.00           Matches: 30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 1                      Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-097-340-153 (1-467)
OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGGAGCGACGACATAAAATCC 60
Db 397 Cysaspleucysaspargillellelleglyaspargglutrpalaahisileysse 416
OY 61 AAATCCCACTTGAAACCACTGAAGAAAGA 90
Db 417 LysSerHisLeuansGlnLeuLysLysArg 426

RESULT 6
US-10-097-340-153
; Sequence 153, Application US/10097340
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xunlei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-153

Alignment Scores:
Pred. No.: 2,17e-15      Length: 467
Score: 164.00           Matches: 30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 24                      Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-097-340-153 (1-467)
OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGGAGCGACGACATAAAATCC 60
Db 397 Cysaspleucysaspargillellelleglyaspargglutrpalaahisileysse 416
OY 61 AAATCCCACTTGAAACCACTGAAGAAAGA 90
```



```

Db          417  LysSerHisLeuAsnGlnLeuIleuLysLysArg 426
|||||
RESULT 7
US-60-389-987-773
; Sequence 773, Application US/60389987
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 6600088.465P2
; CURRENT APPLICATION NUMBER: US/60/389.987
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 773
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-389-987-773

Alignment Scores:
Pred. No.: 2.7          Length: 294
Score: 64.00           Matches: 10
Percent Similarity: 65.52% Conserves: 9
Best Local Similarity: 34.46% Mismatches: 10
Query Match: 39.02%      Indels: 0
DB: 27                  Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-389-987-773 (1-294)
OY 1 TGTGACCTCTGTGATCGAATCATCTATTGGGGATCGCGAATGGCGACGACATATAATCC 60
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 75 CysLysValLysCysLysAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 94
OY 61 AAATCCCACTTGAAACCACTGAAGAAA 87
   ||| ||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 95 LysLysHisAlaAsnLysValLysArg 103

RESULT 8
US-60-412-418-773
; Sequence 773, Application US/60412418
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 6600088.465P3
; CURRENT APPLICATION NUMBER: US/60/412.418
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 773
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-412-418-773

Alignment Scores:
Pred. No.: 2.7          Length: 294
Score: 64.00           Matches: 10
Percent Similarity: 65.52% Conserves: 9
Best Local Similarity: 34.46% Mismatches: 10
Query Match: 39.02%      Indels: 0
DB: 27                  Gaps: 0

```

```

Best Local Similarity: 34.48% Mismatches: 10
Query Match: 39.02% Indels: 0
DB: 27 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-412-418-773 (1-294)

QY      1 TGTGACCTCTGTGATTCGAATCATCATTCATGGCGAGTCCGAATGCCAATGCAGCAGCATTAATCC 60
        ||| ::::||| :||||| |::| :| :| ||||| :|||
Db       75 CysLysValCyScysAlaLeuLeuIleSerGIuSerGIuSlnGlyLeuAlaHisTYrGIuSer 94
QY      61 AAATCCCACCTTGAAACAACACTGAGAATA 87
        ||| ||| ||| ::::::||| ::::
Db       95 LysLYshISaLaAsnLYsValLYsArg 103

RESULT 9
US-60-167-217-10879
; Sequence 10879, Application US/60167217
; GENERAL INFORMATION:
; APPLICANT: LI, Peter W. D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES
; FILE REFERENCE: CL000152
; CURRENT APPLICATION NUMBER: US/60/167,217
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 23195
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10879
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Drosophila
US-60-167-217-10879

Alignment Scores:
Pred. No.:          2.88           Length:         635
Score:             64.00           Matches:         11
Percent Similarity: 57.14%         Conservative:     5
Best Local Similarity: 39.02%       Mismatches:      12
Query Match:       39.02%           Indels:          0
DB:                27              Gaps:            0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-167-217-10879 (1-635)

QY      1 TGTGACCTCTGTGATTCGAATCATCATTCATGGCGAGTCCGAATGCCAATGCAGCAGCATTAATCC 60
        ||| ::||| ||||| |::| :||| :||| ||||| :|||
Db      255 CysGIuLeuCyScysAspValThrCysThrGIyThraspaLatyrAlaAlaHisValArgLy 274
QY      61 AAATCCCACCTTGAAACAACACTGAG 84
        ||| ||| ::|||
Db      275 AlaLYshISaLaAsnValLYs 282

RESULT 10
US-60-173-464-8824
; Sequence 8824, Application US/60173464
; GENERAL INFORMATION:
; APPLICANT: LI, Peter W.D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; FILE REFERENCE: CL000173
; CURRENT APPLICATION NUMBER: US/60/173,464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8824
; LENGTH: 884
; TYPE: PRT
; ORGANISM: Drosophila
US-60-173-464-8824

Alignment Scores:
Pred. No.:          2.97           Length:         884
Score:             64.00           Matches:         11
```

Percent Similarity: 57.14%
 Best Local Similarity: 39.29%
 Query Match: 27
 DB: 27

Conservative: 5
 Mismatches: 12
 Indels: 0
 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-173-464-8824 (1-884)

QY 1 TGTGACCTCTGTGATCGAATCATTCATTGGGATCGCGAATGGCGACGCACATTAATCC 60

DB 255 CysG1bLencysAspValThrCysThrGlyThrAspAlaThrAlaHisValArgGly 274

QY 61 AAATCCACTTGAACCACTGAG 84

DB 275 AlaLysHisGlnAsnValValLys 282

RESULT 11

PCT-US01-18569-3108

; Sequence 3108, Application PC/TUS0118569

; GENERAL INFORMATION:

; APPLICANT: Human Genome Sciences, Inc.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA133PCT

; CURRENT APPLICATION NUMBER: PCT/US01/18569

; PRIOR FILING DATE: 2001-06-07

; PRIOR APPLICATION NUMBER: 60/209,467

; NUMBER OF SEQ ID NOS: 4360

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3108

; LENGTH: 617

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (64)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

PCT-US01-18569-3108

Alignment Scores:

Pred. No.: 4.08
 Score: 63.00
 Percent Similarity: 60.87%
 Best Local Similarity: 39.13%
 Query Match: 38.41%
 DB: 1

Length: 617
 Matches: 9
 Conservative: 5
 Mismatches: 9
 Indels: 0
 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x PCT-US01-18569-3108 (1-617)

QY 1 TGTGACCTCTGTGATCGAATCATTCATTGGGATCGCGAATGGCGACGCACATTAATCC 60

DB 69 CysSerIleCysAspLeuProValHisSerAsnLysGlnThrSerGlnHisIleAsnGly 88

QY 61 AAATCCAC 69

DB 89 AlasErHis 91

RESULT 12

US-10-264-049-3108

; Sequence 3108, Application US/10264049

; GENERAL INFORMATION:

; APPLICANT: Birse et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PA133P1

; CURRENT APPLICATION NUMBER: US/10/264,049

; PRIOR FILING DATE: 2002-10-04

; PRIOR APPLICATION NUMBER: PCT/US01/18569

; PRIOR FILING DATE: 2001-06-07

; PRIOR APPLICATION NUMBER: US 60/209,467

; NUMBER OF SEQ ID NOS: 4360

; SOFTWARE: PatentIn Ver. 3.1

; SEQ ID NO 3108

; LENGTH: 617

TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (64)
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

US-10-264-049-3108

QY 1 TGTGACCTCTGTGATCGAATCATTCATTGGGATCGCGAATGGCGACGCACATTAATCC 60

DB 293 CysSerIleCysAspLeuProValHisSerAsnLysGlnThrSerGlnHisIleAsnGly 312

QY 61 AAATCCAC 69

DB 313 AlasErHis 315

RESULT 13

US-10-205-219-133

; Sequence 133, Application US/10205219

; GENERAL INFORMATION:

; APPLICANT: Warner-Lambert Company

; APPLICANT: Lee, Kevin

; APPLICANT: Dixon, Alistair

; APPLICANT: Brooksbank, Robert

; APPLICANT: Plincock, Robert

; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain

; FILE REFERENCE: WL-A-018200

; CURRENT APPLICATION NUMBER: US/10/205,219

; PRIOR FILING DATE: 2002-07-24

; PRIOR APPLICATION NUMBER: GB 0118354.0

; PRIOR FILING DATE: 2001-07-27

; NUMBER OF SEQ ID NOS: 197

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 133

; LENGTH: 845

; TYPE: PRT

; ORGANISM: Rattus norvegicus

; FEATURE:

; OTHER INFORMATION: Matrin 3

US-10-205-219-133

Alignment Scores:

Pred. No.: 4.18
 Score: 63.00
 Percent Similarity: 60.87%
 Best Local Similarity: 39.13%
 Query Match: 38.41%
 DB: 26

Length: 845
 Matches: 9
 Conservative: 5
 Mismatches: 9
 Indels: 0
 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-205-219-133 (1-845)

QY 1 TGTGACCTCTGTGATCGAATCATTCATTGGGATCGCGAATGGCGACGCACATTAATCC 60

DB 293 CysSerIleCysAspLeuProValHisSerAsnLysGlnThrSerGlnHisIleAsnGly 312

QY 61 AAATCCAC 69

DB 313 AlasErHis 315

RESULT 14

US-09-594-401

; Sequence 401, Application US/09976594

; GENERAL INFORMATION:

```

APPLICANT: Furness, Michael
APPLICANT: Buchbinder, Jenny
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT APPLICATION NUMBER: US/09/976,594
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 401
LENGTH: 847
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 3676741CD1
US-09-976-594-401

```

```

Alignment Scores:
Pred. No.: 4.19 Length: 847
Score: 63.00 Matches: 9
Percent Similarity: 60.87% Conservative: 5
Best Local Similarity: 39.13% Mismatches: 9
Query Match: 38.41% Indels: 0
DB: 23 Gaps: 0

```

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-976-594-401 (1-847)

OY 1 TGTGACCTCTGTGATCGAATCATCTTGGGATCGCGAATGCGCAGCCACATATAATCC 60

DB 293 CysSerIleCysAspLeuProValHisSerAsnLysGluTrpSerGlnHisIleasnGly 312

OY 61 AAATCCAC 69

DB 313 AlaserHis 315

RESULT 15

US-10-153-668-450
Sequence 450, Application US/10153668

GENERAL INFORMATION:

APPLICANT: HONDA, Goichi

APPLICANT: MATSUDA, Akio

APPLICANT: MURAMATSU, Shuji

APPLICANT: ISHIZAWA, Kenya

TITLE OF INVENTION: State Activating Gene

FILE REFERENCE: 1254-0207P

CURRENT APPLICATION NUMBER: US/10/153,668

CURRENT FILING DATE: 2002-05-24

PRIOR APPLICATION NUMBER: US 60/293,172

PRIOR FILING DATE: 2001-05-25

PRIOR APPLICATION NUMBER: US 60/316,031

PRIOR FILING DATE: 2001-08-31

PRIOR APPLICATION NUMBER: US 60/328,403

PRIOR FILING DATE: 2001-10-12

PRIOR APPLICATION NUMBER: JP 2001-157043

PRIOR FILING DATE: 2001-05-25

PRIOR APPLICATION NUMBER: JP 2001-260681

PRIOR FILING DATE: 2001-08-30

PRIOR APPLICATION NUMBER: JP 2001-313175

PRIOR FILING DATE: 2001-10-10

NUMBER OF SEQ ID NOS: 488

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 450

LENGTH: 847

TYPE: PRT

ORGANISM: Homo sapiens

US-10-153-668-450

Alignment Scores:

```

Pred. No.: 4.19 Length: 847
Score: 63.00 Matches: 9
Percent Similarity: 60.87% Conservative: 5

```

```

Best Local Similarity: 39.13% Mismatches: 9
Query Match: 38.41% Indels: 0
DB: 25 Gaps: 0

```

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-153-668-450 (1-847)

OY 1 TGTGACCTCTGTGATCGAATCATCTTGGGATCGCGAATGCGCAGCCACATATAATCC 60

DB 293 CysSerIleCysAspLeuProValHisSerAsnLysGluTrpSerGlnHisIleasnGly 312

OY 61 AAATCCAC 69

DB 313 AlaserHis 315

Search completed: April 21, 2003, 19:12:19
Job time : 15.7893 secs


```
Db 35 LysSerHisLeuAsnGlnLeuLysLysArg 44
RESULT 2
US-60-453-050-7969
; Sequence 7969, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, MAY
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7969
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-7969

Alignment Scores:
Pred. No.: 1.66e-15 Length: 85
Score: 164.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-453-050-7969 (1-85)
QY 1 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCGAATGGCGACCATATAATCC 60
Db 15 CysAspLeuCysAspArgIleIleIleGlyAspArgGluTrpAlaAlaHisIleLysSer 34
QY 61 AAATCCCACTTGAACCACTGACAGAAAAGA 90
Db 35 LysSerHisLeuAsnGlnLeuLysLysArg 44

RESULT 3
US-60-453-135-7967
; Sequence 7967, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7967
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-7967

Alignment Scores:
Pred. No.: 1.73e-15 Length: 221
Score: 164.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-453-135-7967 (1-221)
QY 1 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCGAATGGCGACCATATAATCC 60
Db 151 CysAspLeuCysAspArgIleIleIleGlyAspArgGluTrpAlaAlaHisIleLysSer 170
QY 61 AAATCCCACTTGAACCACTGACAGAAAAGA 90

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-380-731-563 (1-411)
QY 1 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCGAATGGCGACCATATAATCC 60
Db 341 CysAspLeuCysAspArgIleIleIleGlyAspArgGluTrpAlaAlaHisIleLysSer 360
```

```
OY 61 AATCCACTGTGACCACTGAAGAAAGA 90
DB 361 LysSerHisLeuAsnGlnLeuLysArg 370

RESULT 6
US-60-453-135-7968
; Sequence 7968, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7968
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-7968

Alignment Scores:
Pred. No.: 1.79e-15 Length: 475
Score: 164.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-453-135-7968 (1-475)
OY 1 TGTGACCTGTGATGATCATCATCTGGGATCGGATCGGACGACGACATAAATCC 60
DB 405 CysAspLeuCyAspArgIleIleIleGlyAspArgGluTrpAlaIleHisIleLysSer 424

OY 61 AATCCACTGTGACCACTGAAGAAAGA 90
DB 425 LysSerHisLeuAsnGlnLeuLysArg 434

RESULT 7
US-60-453-050-7968
; Sequence 7968, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7968
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-7968

Alignment Scores:
Pred. No.: 1.79e-15 Length: 475
Score: 164.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-453-050-7968 (1-475)
OY 1 TGTGACCTGTGATGATCATCATCTGGGATCGGATCGGACGACGACATAAATCC 60
DB 405 CysAspLeuCyAspArgIleIleIleGlyAspArgGluTrpAlaIleHisIleLysSer 424
```

```
OY 61 AATCCACTGTGACCACTGAAGAAAGA 90
DB 425 LysSerHisLeuAsnGlnLeuLysArg 434

RESULT 8
PCT-US02-41115-54
; Sequence 54, Application PC/TUS0241115
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.
; APPLICANT: BOROWSKY, Mark L.
; APPLICANT: BUREFORD, Neil
; APPLICANT: ELIOTY, Vicki S.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: FORSYTHE, Ian J.
; APPLICANT: GORVAD, Ann E.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: KABLE, Amy E.
; APPLICANT: KHARE, Reena
; APPLICANT: LAL, Preeti G.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: LEE, Soo Yeun
; APPLICANT: LI, Joana X.
; APPLICANT: MARQUIS, Joseph P.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: SPRAGUE, William W.
; APPLICANT: SWARNAKER, Anita
; APPLICANT: TANG, Y. Tom
; APPLICANT: CHAWLA, Nalinder K.
; APPLICANT: WARREN, Bridget A.
; APPLICANT: YUE, Henry
; TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-1338 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/41115
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: US 60/343,004
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/347,633
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US 60/351,749
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/359,498
; PRIOR FILING DATE: 2002-02-22
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PERL Program
; SEQ ID NO 54
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 7505904CD1
PCT-US02-41115-54

Alignment Scores:
Pred. No.: 0.955 Length: 232
Score: 64.00 Matches: 10
Percent Similarity: 65.52% Conservative: 9
Best Local Similarity: 34.48% Mismatches: 10
Query Match: 39.02% Indels: 0
DB: 1 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x PCT-US02-41115-54 (1-232)
OY 1 TGTGACCTGTGATGATCATCATCTGGGATCGGATCGGACGACGACATAAATCC 60
DB 75 CysLysValCysCysAlaLeuLeuIleSerGluSerGlnLysLeuAlaHisIleTyGlnSer 94

OY 61 AATCCACTGTGACCACTGAAGAAAGA 87
```

```
Db 95 LyslyshsAlaAsnLysValIysArg 103
RESULT 9
US-10-218-140-3448
; Sequence 3448, Application US/10218140
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES
; FILE REFERENCE: 15966-543 CON
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US/10/218, 140
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/127,728
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/127,636
; PRIOR FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/127,607
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 6322
; SOFTWARE: Curanator Version 1.0
; SEQ ID NO 3448
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-218-140-3448

Alignment Scores:
Pred. No.: 0.967 Length: 302
Score: 64.00 Matches: 10
Percent Similarity: 65.52% Conservative: 9
Best Local Similarity: 34.48% Mismatches: 10
Query Match: 39.02% Indels: 0
Gaps: 0
DB: 6

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-218-140-3448 (1-302)
QY 1 TGTGACCTCTGTGATCGATCATTCATGGGATCGCGAATGCGGCACATATAATCC 60
Db 83 CyslyshsValIcysCysAlaLeuLeuIleSerGlnLysLeuAlaHisTyrGlnSer 102
QY 61 AAATCCCACTTGAACCACTGAAGAAA 87
Db 103 LyslyshsAlaAsnLysValIysArg 111

RESULT 10
US-60-452-680-16040
; Sequence 16040, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO01450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16040
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(294)
; OTHER INFORMATION: Xaa - Any Amino Acid
US-60-452-680-16040

Alignment Scores:
Pred. No.: 1.36 Length: 294
```

```
Score: 63.00 Matches: 8
Percent Similarity: 60.87% Conservative: 6
Best Local Similarity: 34.78% Mismatches: 9
Query Match: 38.41% Indels: 0
Gaps: 0
DB: 7

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-452-680-16040 (1-294)
QY 1 TGTGACCTCTGTGATCGATCATTCATGGGATCGCGAATGCGGCACATATAATCC 60
Db 231 CysserIleCysAspLysAlaValPheAspLeuLysAspTyrGlnLysValIysGly 250
QY 61 AAATCCCACT 69
Db 251 LysleuHis 253

RESULT 11
PCT-US03-07002-35
; Sequence 35, Application PC/TUS0307002
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; RICHARDSON, Thomas W.;
; APPLICANT: ELLIOTT, Vicki S.; SPRAGUE, William W.;
; APPLICANT: JIANG, Xin; TANG, Y. Tom;
; APPLICANT: ZEBARJADIAN, Yeganeh; BAUGHN, Mariah R.;
; APPLICANT: JACKSON, Alan A.; KABLE, Amy E.;
; APPLICANT: LEE, Ernestine A.; KHAN, Farrah A.;
; APPLICANT: YUE, Henry; FORSYTHE, Ian J.;
; APPLICANT: RAMKUMAR, Jayalaxmi; GRIFFIN, Jennifer A.;
; APPLICANT: LI, Joana X.; MARQUIS, Joseph P.;
; APPLICANT: LEHR-MASON, Patricia M.; JIN, Pei;
; APPLICANT: HARKINS, Phillips R.; WILSON, Amy D.;
; APPLICANT: SWARNAKAR, Anita; HE, Ann;
; APPLICANT: HAFALIA, April J.A.; TRAN, Bao;
; APPLICANT: DUGGAN, Brendan M.; EMERLING, Brooke M.;
; APPLICANT: BOROWSKY, Mark L.; YAO, Monique G.;
; APPLICANT: CHAWLA, Narinder K.; BURFORD, Neil;
; APPLICANT: KHARE, Reena; LEE, Sally;
; APPLICANT: BECHA, Shanya D.; LEE, Soo Yeun
; TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-1377 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/07002
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: US 60/362,329
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/364,438
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/373,891
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PERL Program
; SEQ ID NO 35
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7509991CD1
PCT-US03-07002-35

Alignment Scores:
Pred. No.: 1.36 Length: 339
Score: 63.00 Matches: 9
Percent Similarity: 60.87% Conservative: 5
Best Local Similarity: 39.13% Mismatches: 9
Query Match: 38.41% Indels: 0
Gaps: 0
DB: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x PCT-US03-07002-35 (1-339)
QY 1 TGTGACCTCTGTGATCGATCATTCATGGGATCGCGAATGCGGCACATATAATCC 60
Db 293 CysserIleCysAspLeuProValHisSerAsnLysGluTyrSerGlnHisIleAsnGly 312
```

[illegible]

```

OY      61  AATGCCAC 69
      |||||
Db      313  AlaserHis 315

RESULT 14
US-09-949-016-11423
/ Sequence 11423, Application US/09949016
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: C1001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ. ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 11423
/ LENGTH: 853
/ TYPE: PRT
/ ORGANISM: Human
US-09-949-016-11423

Alignment Scores:
Pred. No.:      1.42      Length:      853
Score:          63.00     Matches:      9
Percent Similarity: 60.87% Conservative: 5
Best Local Similarity: 39.13% Mismatches:  9
Query Match:     38.41%  Indels:      0
Db:              5       Gaps:      0

US-09-513-151-3_COPY_1121.1210 (1-90) x US-09-949-016-11423 (1-853)
OY      1  TGTGACCGTGTGATCGAATCATTCATGGGATCGGATGGGACGCATTAATCC 60
      ||| :|||:||||| :||| :||| :|||:|||||:|||||
Db      299  CysserillecysaspleuProvalHisserasnlvsglurpserclnHisIleasnslg 318
OY      61  AATGCCAC 69
      |||||
Db      319  AlaserHis 321

RESULT 15
US-09-675-784A-9300
/ Sequence 9300, Application US/09675784A
/ GENERAL INFORMATION:
/ APPLICANT: HARE, ROBERTA S.
/ APPLICANT: SHAW, KAREN J.
/ APPLICANT: SHIMER JR., GEORGE H.
/ APPLICANT: KESSLER, MARCO
/ APPLICANT: NOLLING, JORK
/ APPLICANT: ZENG, QIANDONG
/ APPLICANT: GREENE, JONATHAN R.
/ TITLE OF INVENTION: ASPERGILLUS FUNIGATUS NUCLEIC ACIDS AND POLYPEPTIDES,
/ FILE REFERENCE: 2976-4020US1
/ CURRENT APPLICATION NUMBER: US/09/675,784A
/ CURRENT FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/156,338
/ PRIOR FILING DATE: 1999-09-29
/ NUMBER OF SEQ. ID NOS: 13925
/ SEQ ID NO 9300
/ LENGTH: 242
/ TYPE: PRT
/ ORGANISM: Aspergillus fumigatus
US-09-675-784A-9300

Alignment Scores:

```


Pred. No.:	10.3	length:	22.2
Score:	57.00	Matches:	7
Percent Similarity:	53.57%	Conservative:	8
Best Local Similarity:	25.00%	Mismatches:	13
Query Match:	34.76%	Indels:	0
DB:	5	Gaps:	0

US-09-513-151-3_COPY_1121_1210 (1-90) X US-09-675-784A-9300 (1-242)

Oy 1 TGTGACCCCTTGATGAGCATCATATTGGGAGTGCAGCAATGAGCGCACATAAATCC 60
 |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
 Db 207 CysctutleCysserArgThMetAlathrcInsprlntrpAsrIlehtIsLeauAnGly 226
 Oy 61 AAATGCCACTTGAAACAACACTGAAG 84
 :::::|||:::|||
 Db 227 ArgAlahIslvsArgAlalaleLys 234
 ::::||||

Search completed: April 21, 2003, 18:58:58
Job time : 4.07555 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 21, 2003, 18:49:25 ; Search time 2.57626 Seconds
(without alignments)
5283.055 Million cell updates/sec

Title: US-09-513-151-3_COPY_1121_1210

Perfect score: 164

Sequence: 1 TGTGACCTCTGTGATCGAAT.....TGACCACTGAAGAAGAAGA 90

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 288829 seqs, 75613885 residues

Total number of hits satisfying chosen parameters: 577658

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV-xlh
-O-/cgn2_1/USPTO_spool/US09513151/runat_15042003_141145_26460/app_query.fasta_1.2446
-DB-Published Applications_AA -GWT-fastan -SUFFIX-n2p.rapb -MINMATCH-0.1
-LOOPEXT-0 -LOOPEXT-0 -UNITS-bits -SMART-1 -END-1 -MATRIX-biosum62
-TRANS-human40.cdi -LIST-45 -DOCALLIGN-200 -THR_SCORE-pct -THR_MAX-100
-THR_MIN-0 -ALIGN-15 -MODE-LOCAL -OUTFWT-pco -NORM-ext -HEAPSIZE-500 -MINLEN-0
-MAXLEN-2000000000 -USER-US09513151_@CGN_1.1.15_runat_15042003_141145_26460
-NCPU-6 -ICPU-3 -NO_XLPEXY -NO_MMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database: Published Applications_AA:

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	34.8	737	9	US-10-102-806-731
2	54	32.9	123	9	US-09-986-480-246
3	54	32.9	187	9	US-09-986-480-368
4	53.5	32.6	238	10	US-09-764-864-901

C	5	51	31.5	95	10	US-09-731-872-320	Sequence 320, App
	6	51	31.1	206	10	US-09-764-864-1012	Sequence 1012, App
	7	51	31.1	206	10	US-09-764-864-1441	Sequence 1441, App
	8	50	30.5	166	9	US-10-102-806-571	Sequence 571, App
	9	50	30.5	179	10	US-09-925-302-679	Sequence 679, App
	10	50	30.5	229	10	US-09-866-562-92	Sequence 92, App
C	11	49.5	30.6	578	9	US-09-999-248-2	Sequence 2, App
	12	48.5	29.6	762	9	US-10-101-464-114	Sequence 114, App
	13	48.5	29.6	809	9	US-10-105-695-2	Sequence 2, App
	14	48.5	29.6	809	10	US-10-105-694-2	Sequence 2, App
	15	48.5	29.6	809	10	US-09-747-521-2	Sequence 2, App
	16	48.5	29.6	809	12	US-10-106-014-2	Sequence 2, App
C	17	48	29.6	52	9	US-09-941-831-29	Sequence 29, App
	18	48	29.6	53	10	US-09-864-761-37524	Sequence 43910, App
	19	48	29.6	74	10	US-09-864-761-37524	Sequence 37524, App
C	20	48	29.6	192	9	US-09-764-868-1192	Sequence 1192, App
	21	48	29.3	221	9	US-09-738-626-4142	Sequence 4142, App
C	22	48	29.6	322	9	US-09-828-523-10	Sequence 10, App
C	23	48	29.6	322	9	US-09-828-523-12	Sequence 12, App
C	24	48	29.6	331	9	US-09-828-523-66	Sequence 66, App
C	25	48	29.6	331	9	US-09-828-523-70	Sequence 70, App
C	26	48	29.3	426	9	US-09-966-781-2	Sequence 2, App
C	27	48	29.6	687	10	US-09-965-602-12	Sequence 12, App
C	28	48	29.6	821	9	US-09-764-868-883	Sequence 883, App
C	29	48	29.6	878	12	US-10-060-332-2	Sequence 2, App
C	30	47	29.0	208	10	US-09-764-864-833	Sequence 833, App
C	31	47	29.0	208	10	US-09-764-877-1292	Sequence 1292, App
C	32	47	28.7	500	9	US-10-124-800-30	Sequence 30, App
C	33	47	28.7	753	10	US-09-908-180-2	Sequence 2, App
C	34	47	29.0	807	10	US-09-847-046-2	Sequence 2, App
C	35	47	28.7	1185	10	US-09-815-242-11466	Sequence 11466, App
	36	47	28.7	1503	9	US-10-124-800-6	Sequence 6, App
	37	47	28.7	1503	9	US-10-124-800-28	Sequence 28, App
	38	46.5	28.4	299	10	US-09-827-854-2	Sequence 2, App
	39	46.5	28.4	299	10	US-09-827-854-3	Sequence 3, App
	40	46.5	28.4	317	9	US-09-870-759-130	Sequence 130, App
	41	46.5	28.4	317	9	US-09-802-640-18	Sequence 18, App
	42	46.5	28.4	317	10	US-09-827-854-15	Sequence 15, App
	43	46.5	28.4	317	10	US-09-827-854-16	Sequence 16, App
	44	46.5	28.4	498	9	US-10-037-667-5	Sequence 5, App
	45	46	28.0	31	10	US-09-938-315-44	Sequence 44, App

ALIGNMENTS

RESULT 1
US-10-102-806-731
Sequence 731, Application US/10102806
Publication No. US20030054421A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA103PIC1
CURRENT APPLICATION NUMBER: US/10/102,806
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/925,298
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05881
PRIOR FILING DATE: 2000-03-08
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 846
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 731
LENGTH: 737
TYPE: PRT
ORGANISM: Homo sapiens
US-10-102-806-731
Alignment Scores:
Pred. No.: 3.53
Score: 57.00
Percent Similarity: 53.57%
Length: 737
Matches: 11
Conservative: 4

```

Best Local Similarity: 39.29%      Mismatches: 13
Query Match:          34.76%     Indels:       0
DB:                   9           Gaps:        0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-102-806-731 (1-737)

OY    1 TGTGACCTTGATGCAGCAATCATCTTTGGGATCGCGAATTGGCGACGCCACATAAATCC   60
      |||:::||||| | | | | | | | | | | | | :|||::|
Db    47 CysgIstleuCysASpsvalSerCysthrGIyAlaaSPalAtyTfAlAAIAHisIleaRgily 66
      |||               ::|||

OY    61 AAATCCCACTTCAGAACCAACTGAGA 84
      |||                ::|||
Db    67 AlalyshISGLnlySvalVallys 74

RESULT 2
US-09-986-480-246
; Sequence 246, Application US/09986480
; Publication No. US20030027999A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 143 Human Secreted Proteins
FILE REFERENCE: PS500P1
CURRENT APPLICATION NUMBER: US/09/986,480
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: PCT/US00/12788
PRIOR FILING DATE: 2000-05-11
PRIOR FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 456
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 246
LENGTH: 123
TYPE: PRF
ORGANISM: Homo sapiens
US-09-986-480-246

Alignment Scores:
Pred. NO.:         7.92             Length:      123
Score:            54.00            Matches:      12
Percent Similarity: 43.18%          Conservative: 7
Best local Similarity: 32.27%       Mismatches:   11
Query Match:      32.93%            Indels:       14
DB:              9                 Gaps:          1

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-986-480-246 (1-123)

OY    1 TGTGACCTTGATGCAGCAATCATCTTTGGGATCGCGAATGGCGACGCCACATAAATCC   60
      |||:::||||| | | | | | | | | | | | | :|||::|
Db    6 CysasvnaIcysAspcysvalVallysaspSerILeaSnPhelenaSpHsisIleaSn gly 25
      |||               ::|||

OY    61 AAATCCCAC-----TTGAACCAA 78
      |||                  |||
Db    26 LysLysHSIGlnArgASnleucIlmetSetMetArGaIgluarGerThrLeuaSpGln 45
      ::|||::|

OY    79 CTGAAGAAAAGA 90
      ::|||::|
Db    46 ValLySLysarg 49

RESULT 3
US-09-986-480-368
; Sequence 368, Application US/09986480
; Publication No. US20030027999A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 143 Human Secreted Proteins
FILE REFERENCE: PS500P1
CURRENT APPLICATION NUMBER: US/09/986,480
CURRENT FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: PCT/US00/12788
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: US 60/134,068
PRIOR FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 456
```

```

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 368
LENGTH: 187
TYPE: PRT
ORGANISM: Homo sapiens
US-09-986-480-368

Alignment Scores:
Pred. No.:      8.38      Length:      187
Score:          54.00     Matches:       12
Percent Similarity: 43.18% Conservative:   7
Best Local Similarity: 27.27% Mismatches:    11
Query Match:      32.93% Indels:         14
DB:              Gaps:           1

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-986-480-368 (1-187)
QY      1 TGTGACCTCTGTGATCGAATCATCATTTGGGGATCCGCAGATGGGCCCAATAAAATCC 60
        |||:::|||||||:::|||||
Db      70 CysAsnValCysAspCysValValLysAspSerIleasnPheLeuaspHisIleasncl 89
        ||| |||
QY      61 AAATCCCAC-----TTGACCAC 78
        ||| |||
Db      90 LysLysHisSlnArgsnLenuglMetserMetargValgluArgSerThrLeuaspeln 109
        :::|||||
QY      79 CTGAAGAANAAGA 90
Db      110 ValLysLysArg 113

RESULT 4
US-09-764-864-901
Sequence 901, Application US/09764864
Patent No. US20020132753A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT223
CURRENT APPLICATION NUMBER: US/09/764,864
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 901
LENGTH: 238
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (135)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (153)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (156)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (185)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (195)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (196)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (210)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-901

Alignment Scores:
Pred. No.:      10.3     Length:      238
Score:          53.50     Matches:       13

```

```

; NUMBER OF SEQ ID NOS: 1792
;
; SOFTWARE: PatentIn Ver. 2.0

```

; PRIOR APPLICATION NUMBER: 09/
 ; PRIOR FILING DATE: 2001-08-10

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-925-302-679 (1-179)

ORGANISM: *Homo sapiens*

US-09-999-248-2

Alignment Scores:

Pred. No.:	47.4	Length:	578
Score:	49.50	Matches:	8
Percent Similarity:	50.00%	Conservative:	5
Best Local Similarity:	30.77%	Mismatches:	6
Query Match:	30.56%	Indels:	7
DB:	9	Gaps:	1

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-999-248-2 (1-578)

QY 89 CTTTCTTCATGTCG-----TTCAAGTGGCATTTGATTGATTTT 51

Db 457 LeuTyRphelIleTyRyAlCysArgAspIleGlnSerPheArgTrpPheAlaAspLeu 476

QY 50 TGGCGTGGCCCATTCGCGCA 33

Db 477 CysMetLeuHisAsnLys 482

RESULT 12

US-10-101-464A-114

; Sequence 114, Application US/10101464A

; Publication No. US20030046728A1

; GENERAL INFORMATION:

; APPLICANT: Strabala, Timothy

; APPLICANT: Nieuwenhuizen, Nicolaas

; APPLICANT: Higgins, Colleen M.

; TITLE OF INVENTION: Compositions Isolated from Plant Cells

; TITLE OF INVENTION: and their use in the Modification of Plant Cell Signaling

; FILE REFERENCE: 11000.1020C2

; CURRENT APPLICATION NUMBER: US/10/101,464A

; PRIOR FILING DATE: 2002-03-18

; PRIOR APPLICATION NUMBER: 09/704,302

; PRIOR FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: 09/228,986

; PRIOR FILING DATE: 1999-01-12

; PRIOR APPLICATION NUMBER: 60/162,866

; PRIOR FILING DATE: 1999-11-01

; PRIOR APPLICATION NUMBER: PCT/US00/00724

; PRIOR FILING DATE: 2000-01-11

; NUMBER OF SEQ ID NOS: 989

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 114

; LENGTH: 762

; TYPE: PRT

; ORGANISM: Pinus radiata

US-10-101-464A-114

Alignment Scores:

Pred. No.:	69.9	Length:	762
Score:	48.50	Matches:	10
Percent Similarity:	63.16%	Conservative:	2
Best Local Similarity:	52.63%	Mismatches:	6
Query Match:	29.57%	Indels:	1
DB:	9	Gaps:	1

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-101-464A-114 (1-762)

QY 15 TCGAATCATTCATTCGGGATCGCGATGCGACATTAATCCAAATCCACTT 71

Db 261 SerAsnHisHisTrp---SerLeuAlaGlyProValAlaAlaIleArgValProLeu 278

RESULT 13

US-10-105-695-2

; Sequence 2, Application US/10105695

; Publication No. US20020197272A1

; GENERAL INFORMATION:

; APPLICANT: Galloway, Darrel R.

; APPLICANT: Mateczun, Alfred J.

; TITLE OF INVENTION: Methods for Protection Against Lethal Infection with Bacillus Ant

; FILE REFERENCE: 22727/04115

; CURRENT APPLICATION NUMBER: US/10/105,695

; CURRENT FILING DATE: 2002-03-25

; PRIOR APPLICATION NUMBER: US 09/747,521

; PRIOR FILING DATE: 2000-12-21

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 809

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-10-105-695-2

Alignment Scores:

Pred. No.:	70.4	Length:	809
Score:	48.50	Matches:	9
Percent Similarity:	53.85%	Conservative:	5
Best Local Similarity:	34.62%	Mismatches:	7
Query Match:	29.57%	Indels:	5
DB:	9	Gaps:	1

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-105-695-2 (1-809)

QY 24 CATTCGGGATTCGCGAATG-----GGCAGCGCACATTAATCCAAATCCCA 68

Db 313 HistTrpSerAspSerLeuSerGluGluGlyArgGlyLeuLeuLysLysLeuGlnIlePro 332

QY 69 CTGGAACCACTGAGAA 86

Db 333 IleGluProLysLysAsp 338

RESULT 14

US-10-105-694-2

; Sequence 2, Application US/10105694

; Publication No. US20030003109A1

; GENERAL INFORMATION:

; APPLICANT: Mateczun, Alfred J.

; APPLICANT: Galloway, Darrel R.

; TITLE OF INVENTION: Methods for Protection Against Lethal Infection with Bacill

; FILE REFERENCE: 22727/04116

; CURRENT APPLICATION NUMBER: US/10/105,694

; PRIOR FILING DATE: 2002-03-25

; PRIOR APPLICATION NUMBER: US 09/747,521

; PRIOR FILING DATE: 2000-12-21

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 809

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-10-105-694-2

Alignment Scores:

Pred. No.:	70.4	Length:	809
Score:	48.50	Matches:	9
Percent Similarity:	53.85%	Conservative:	5
Best Local Similarity:	34.62%	Mismatches:	7
Query Match:	29.57%	Indels:	5
DB:	9	Gaps:	1

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-105-694-2 (1-809)

QY 24 CATTCGGGATTCGCGAATG-----GGCAGCGCACATTAATCCAAATCCCA 68

Db 313 HistTrpSerAspSerLeuSerGluGluGlyArgGlyLeuLeuLysLysLeuGlnIlePro 332

QY 69 CTGGAACCACTGAGAA 86

Db 333 IleGluProLysLysAsp 338

RESULT 15

US-09-747-521-2

; Sequence 2, Application US/09747521

; Patent No. US20020051791A1

; GENERAL INFORMATION:

APPLICANT: Galloway, Darrel
APPLICANT: Mateczun, Alfred
TITLE OF INVENTION: Method for Protection Against Lethal Infection with Bacillus Ant
FILE REFERENCE: 22727/04079
CURRENT APPLICATION NUMBER: US/09/747,521
CURRENT FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 2
LENGTH: 809
TYPE: PRF
ORGANISM: Bacillus anthracis
US-09-747-521-2

Alignment Scores:

Pred. No.:	70.4	Length:	809
Score:	48.50	Matches:	9
Percent Similarity:	53.858	Conservative:	5
Best Local Similarity:	34.628	Mismatches:	7
Query Match:	29.574	Indels:	5
DB:	10	Gaps:	1

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-747-521-2 (1-809)

QY	24	CATTGGGATCGCGAATG-----GGCAGCGCACATAAATCCAATCCA	68
DB	313	HistpSerAspserLeuSerGluGluGlyArgGlyLeuLeuLysLysLeuGlnIlePro	332
QY	69	CTTGACCAACTGAGAA	86
DB	333	IleGluProLysLysAsp	338

Search completed: April 21, 2003, 19:01:37
Job time : 2.57626 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 21, 2003, 18:41:35 ; Search time 0.971375 Seconds
(without alignments)
5452.195 Million cell updates/sec

Title: US-09-513-151-3_COPY_1121_1210

Perfect score: 1 TGTGACCTCTGTGATCGAAT.....TGAACCACTGAGAGAAAAGA 90

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-O=/cgn2_1/USPRO.spool/US09513151/runat_15042003_141144_26406/app.query.fasta_1.2446
-DB=Issued_Patents_AA -OPMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -NCPU=6 -ICPU=3
-USER=US09513151_ECGN_1.1_23_runat_15042003_141144_26406 -NCPU=6 -ICPU=3
-NO.XLPYX -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELext=7

Database :

Issued_Patents_AA:*
1: /cgn2_6/plodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/plodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/plodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/plodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/plodata/1/1aa/PCUTUS.COMB.pep:*
6: /cgn2_6/plodata/1/1aa/Backfill1est.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	32.9	500	3	US-09-141-000-2
2	53	32.3	418	3	US-09-141-000-6
3	53	32.3	431	2	US-08-836-620A-19
4	53	32.3	435	3	US-09-040-508-2
5	53	32.3	435	4	US-09-500-654-2
6	53	32.3	438	3	US-09-141-000-4
7	51	31.1	653	4	US-09-443-184-50
8	49	30.2	136	4	US-09-288-143-190
9	49	30.2	226	4	US-08-679-493A-23
10	48.5	29.6	762	4	US-09-228-986-114
11	48.5	29.6	776	1	US-08-021-601-2
12	48.5	29.6	776	1	US-08-082-849B-2

13	48.5	29.6	776	5	PCT-US94-01624-2	Sequence 2, Appl1
14	48	29.6	878	4	US-09-735-934A-2	Sequence 2, Appl1
15	47.5	29.0	244	3	US-09-135-782-4	Sequence 4, Appl1
16	47.5	29.0	244	4	US-09-193-191-4	Sequence 4, Appl1
17	46.5	28.4	220	2	US-08-726-306A-29	Sequence 29, Appl1
18	46.5	28.4	317	1	US-07-709-949-2	Sequence 2, Appl1
19	46	28.0	31	4	US-08-602-999A-44	Sequence 44, Appl1
20	46	28.0	31	4	US-08-278-865-44	Sequence 44, Appl1
21	46	28.0	31	4	US-09-500-114-44	Sequence 38, Appl1
22	46	28.0	170	4	US-08-444-818-38	Sequence 4, Appl1
23	46	28.4	216	2	US-08-737-716-4	Sequence 2, Appl1
24	46	28.4	382	4	US-08-969-815-2	Sequence 2, Appl1
25	46	28.4	382	4	US-09-120-025-2	Sequence 2, Appl1
26	46	28.4	382	4	US-09-710-481-2	Sequence 2, Appl1
27	46	28.0	418	4	US-09-202-893B-4	Sequence 2, Appl1
28	46	28.0	437	2	US-08-737-716-2	Sequence 4, Appl1
29	46	28.0	502	4	US-09-330-970-1	Sequence 1, Appl1
30	46	28.4	516	4	US-08-867-611-6	Sequence 6, Appl1
31	46	28.4	516	5	PCT-US92-06965A-11	Sequence 11, Appl1
32	46	28.0	518	2	US-08-836-620A-18	Sequence 18, Appl1
33	46	28.0	568	1	US-08-262-424-7	Sequence 7, Appl1
34	46	28.0	568	2	US-08-493-187-7	Sequence 7, Appl1
35	46	28.0	568	2	US-08-717-587A-4	Sequence 4, Appl1
36	46	28.0	568	3	US-08-883-610A-4	Sequence 4, Appl1
37	46	28.0	568	4	US-08-936-094A-4	Sequence 4, Appl1
38	46	28.0	568	5	PCT-US95-07844-7	Sequence 7, Appl1
39	46	28.4	798	4	US-08-867-611-36	Sequence 36, Appl1
40	46	28.4	894	4	US-09-735-934A-4	Sequence 4, Appl1
41	46	28.4	1011	4	US-08-850-328-5	Sequence 5, Appl1
42	46	28.4	1786	4	US-08-444-818-54	Sequence 54, Appl1
43	46	28.4	2261	4	US-08-444-818-66	Sequence 66, Appl1
44	46	28.4	2436	4	US-08-444-818-75	Sequence 75, Appl1
45	46	28.4	2772	4	US-08-444-818-89	Sequence 89, Appl1

ALIGNMENTS

RESULT 1
US-09-141-000-2
US-09-141-000-2 Application US/09141000
Patent No. 6054295
GENERAL INFORMATION:
APPLICANT: Chen, Pang
TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
FILE REFERENCE: 199999Y
CURRENT APPLICATION NUMBER: US/09/141,000
CURRENT FILING DATE: 1998-08-26
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 500
TYPE: PRT
ORGANISM: Human
US-09-141-000-2

Alignment Scores:
Pred. No.: 2.4
Score: 54.00
Percent Similarity: 57.14%
Best Local Similarity: 50.00%
Query Match: 32.93%
DB: 3
Gaps: 2

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-141-000-2 (1-500)

QY	1	TGTGACCTCTGTGATCGA-----ATCACCATTGGGAGATCGGATGCGCAGCGAC	51
DB	244	CysAspPLeuAlaAspArgGluLeuValIleIleGly-----TrrAlaYshIs	260
QY	52	ATTAATCAATCCATCCACTGAGC	75
DB	261	IleProGlyPheSerSerLeuSer	268


```
RESULT 2
US-09-141-000-6
; Sequence 6, Application US/09141000
; Patent No. 6054295
; GENERAL INFORMATION:
; APPLICANT: Chen, Fang
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
; FILE REFERENCE: 199999Y
; CURRENT APPLICATION NUMBER: US/09/141,000
; CURRENT FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Human
US-09-141-000-6

Alignment Scores:
Pred. No.: 3.34 Length: 418
Score: 53.00 Matches: 14
Percent Similarity: 57.14% Conservative: 2
Best Local Similarity: 50.00% Mismatches: 6
Query Match: 32.32% Indels: 6
DB: 3 Gaps: 2

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-141-000-6 (1-418)

QY 1 TGTGACCTCTGTGATCGA-----ATCATCATTTGGGAGATCGCAATGGCAGCGCAC 51
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 269 CysAspLeuAlaAspArgGluLeuValIleIleGly-----TrpAlaLysHis 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 52 ATAAATCCAAATCCCACTTGAAAC 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 286 IleProGlyPheSerThrLeuSer 293

RESULT 3
US-08-836-620A-19
; Sequence 19, Application US/08836620A
; Patent No. 5958710
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Orphan receptor
; NUMBER OF SEQUENCES: 19
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,620A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/03933
; FILING DATE:
; APPLICATION NUMBER: GB 9518272.1
; FILING DATE: 08-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9605550.4
; FILING DATE: 15-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9607532.0
; FILING DATE: 11-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9609576.5
; FILING DATE: 08-MAY-1996
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

```
ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-836-620A-19

Alignment Scores:
Pred. No.: 3.36 Length: 431
Score: 53.00 Matches: 11
Percent Similarity: 58.33% Conservative: 3
Best Local Similarity: 45.83% Mismatches: 10
Query Match: 32.32% Indels: 0
DB: 2 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-08-836-620A-19 (1-431)

QY 1 TGTGACCTCTGTGATCGATCATCATTTGGGAGATCGCAATGGCAGCGCAATAATCC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 244 CysAspLeuAlaAspArgGluLeuValIleIleSerTrpAlaLysHisIleProGly 263
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 AAATCCCACTTG 72
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 264 PheSerAsnLeu 267

RESULT 4
US-09-040-508-2
; Sequence 2, Application US/09040508
; Patent No. 6069239
; GENERAL INFORMATION:
; APPLICANT: Mathias, Stephen L.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,508
; FILING DATE: 17-MAR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9705451.4
; FILING DATE: 17-MAR-1997
; APPLICATION NUMBER: GB 9803289.9
; FILING DATE: 16-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-30008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-040-508-2

Alignment Scores:
Pred. No.: 3.37 Length: 435
Score: 53.00 Matches: 14
Percent Similarity: 57.14% Conservative: 2
Best Local Similarity: 50.00% Mismatches: 6
```

```
Query Match: 32.32% Indels: 6
DB: 3 Gaps: 2
US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-040-508-2 (1-435)
QY 1 TGTGACCTCTGTGATCGA-----ATCATCATTTGGGATCGCGAATGGCAGCGCAC 51
DB 246 CysAspLeuAlaAspArgGluLeuValIleIleGly-----TrrPalalysHs 262

QY 52 ATAAATCCAAATCCCACTTGAC 75
DB 263 IleProGlyPheSerThrLeuSer 270

RESULT 5
US-09-500-654-2
; Sequence 2, Application US/09500654
; Patent No. 6359116
; GENERAL INFORMATION:
; APPLICANT: Stephen L. Mathias
; FILE REFERENCE: GP-30008-D1
; CURRENT APPLICATION NUMBER: US/09/500,654
; EARLIER FILING DATE: 2000-02-09
; EARLIER APPLICATION NUMBER: 9705451.4
; EARLIER FILING DATE: 1997-03-17
; EARLIER APPLICATION NUMBER: 9803289.9
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 435
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-500-654-2

Alignment Scores:
Pred. No.: 3.37 Length: 435
Score: 53.00 Matches: 14
Percent Similarity: 57.14% Conservative: 2
Best Local Similarity: 50.00% Mismatches: 6
Query Match: 32.32% Indels: 6
DB: 4 Gaps: 2
US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-500-654-2 (1-435)
QY 1 TGTGACCTCTGTGATCGA-----ATCATCATTTGGGATCGCGAATGGCAGCGCAC 51
DB 246 CysAspLeuAlaAspArgGluLeuValIleIleGly-----TrrPalalysHs 262

QY 52 ATAAATCCAAATCCCACTTGAC 75
DB 263 IleProGlyPheSerThrLeuSer 270

RESULT 6
US-09-141-000-4
; Sequence 4, Application US/09141000
; Patent No. 6054295
; GENERAL INFORMATION:
; APPLICANT: Chen, Fang
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
; FILE REFERENCE: 19999Y
; CURRENT APPLICATION NUMBER: US/09/141,000
; EARLIER FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Human
US-09-141-000-4

Alignment Scores:
Pred. No.: 3.4 Length: 458
Score: 53.00 Matches: 14
Percent Similarity: 57.14% Conservative: 2
Best Local Similarity: 50.00% Mismatches: 6
Query Match: 32.32% Indels: 6
DB: 3 Gaps: 2
US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-141-000-4 (1-458)
QY 1 TGTGACCTCTGTGATCGA-----ATCATCATTTGGGATCGCGAATGGCAGCGCAC 51
DB 269 CysAspLeuAlaAspArgGluLeuValIleIleGly-----TrrPalalysHs 285

QY 52 ATAAATCCAAATCCCACTTGAC 75
DB 286 IleProGlyPheSerThrLeuSer 293

RESULT 7
US-09-443-184-50
; Sequence 50, Application US/09443184A
; Patent No. 6372431
; GENERAL INFORMATION:
; APPLICANT: Cunningham, Mary Jane
; APPLICANT: Zweiger, Gary
; APPLICANT: Kaser, Matthew R.
; APPLICANT: Panzer, Scott
; APPLICANT: Yue, Henry
; APPLICANT: Selhammer, Baughn, Mariah
; APPLICANT: Azimzal, Yalda
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: MAMMALIAN TOXICOLOGICAL RESPONSE MARKERS
; FILE REFERENCE: PC-0007 US
; CURRENT APPLICATION NUMBER: US/09/443,184A
; EARLIER FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 50
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6372431 3511087CD1
US-09-443-184-50

Alignment Scores:
Pred. No.: 7.57 Length: 653
Score: 51.00 Matches: 7
Percent Similarity: 56.52% Conservative: 6
Best Local Similarity: 30.43% Mismatches: 10
Query Match: 31.10% Indels: 0
DB: 4 Gaps: 0
US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-443-184-50 (1-653)
QY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCAGCGCACATAAATCC 60
DB 541 CysTrpMetCysGlyAsnGlyAsnGlySerGluLeuValIleIleGly-----TrrPalalysHs 560

QY 61 AAATCCGAC 69
DB 561 GluLysHs 563

RESULT 8
US-09-288-143-190
; Sequence 190, Application US/09288143
; Patent No. 6433139
; GENERAL INFORMATION:
; APPLICANT: Brewer et al.
; TITLE OF INVENTION: 53 Human Secreted Proteins
; FILE REFERENCE: P2018P1
; CURRENT APPLICATION NUMBER: US/09/288,143
```

```

; SEQ ID NO 23
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(226)
; OTHER INFORMATION: X is selenocysteine.
US-08-679-493A-23

Alignment Scores:
Pred. No.: 12.7 Length: 226
Score: 49.00 Matches: 8
Percent Similarity: 50.00% Conservative: 2
Best Local Similarity: 40.00% Mismatches: 10
Query Match: 30.25% Indels: 0
DB: 4 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-08-679-493A-23 (1-226)
QY 86 TTCCTCAGTTGGTTCAGTGGGATTTGGATTTATGTCGCTGCCATTCGCATCCCA 27
Db 87 TrrPhcCystTrpPhcProrTrpGluTrpProSerThcCysAlaAlaGlyGlyGluPro 106
RESULT 10
US-09-228-986-114
; Sequence 114, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Muewenhutzen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and their use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228, 986
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 114
; LENGTH: 762
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-114

Alignment Scores:
Pred. No.: 19.5 Length: 762
Score: 48.50 Matches: 10
Percent Similarity: 63.16% Conservative: 2
Best Local Similarity: 52.63% Mismatches: 6
Query Match: 29.57% Indels: 1
DB: 4 Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-228-986-114 (1-762)
QY 15 TCGATTCATCATGTGGGATCGCAATGGCAGCGCACATATAATCCAAATCCCACTT 71
Db 261 SerAsnHisHisStrp---SerLeuAlaGlyProValAlaAlaIleArgValProLeu 278
RESULT 11
US-08-021-601-2
; Sequence 2, Application US/08021601
; Patent No. 5591631
; GENERAL INFORMATION:
; APPLICANT: Leppia, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Nichols, Peter J.
; APPLICANT: Atora, Naveen
; APPLICANT: Singh, Yogendra
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Needle & Rosenberg, P. C.

```

```

STREET: 133 Carnegie Way, Suite 400
City: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,601
FILING DATE: 19930212
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Spratt, Gwendolyn D.,
REGISTRATION NUMBER: 36,016
REFERENCE/DOCKET NUMBER: 1414.057
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9680
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 776 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-021-601-2

Alignment Scores:
Pred. No.:          Length:      776
Score:              Matches:     9
Percent Similarity: 48.50%
                    Conservative: 5
Best Local Similarity: 53.85%
                    Mismatches:   7
Query Match:        29.57%       Indels:    5
                                   Caps:      1

US-09-513-151-3-COPY_1121_1210 (1-90) x US-08-021-601-2 (1-776)
QY      24 CATTGGGAGTCGCAGATG-----GGCAGCGCACATAAATCCAAATCCA 68
         |||||           :::            |||             ||:::|||||
Db      280 Histrpsrarpserleusercjgluglaryargelyleuylstysleughnlltpro 299
         ::|||         :|||||
QY      69 CTGAACCACTGAAGA 86
         ::|||         :|||||
Db      300 IlegluProlystLysasp 305

RESULT 12
US-08-082-849B-2
; Sequence 2, Application US/08082849B
; Patent No. 5677274
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Kimpel, Kurt R.
; APPLICANT: Alora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
; TITLE OF INVENTION: Related Methods
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

      APPLICATION NUMBER: US/08/082,849B
      FILING DATE: 25-JUN-1993
      CLASSIFICATION: 514
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 08/021,601
      FILING DATE: 12-FEB-1993
      ATTORNEY/AGENT INFORMATION:
      NAME: Weber, Kenneth A.
      REGISTRATION NUMBER: 31,677
      REFERENCE/DOCKET NUMBER: 15280-161-1
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (415) 576-0300
      TELEFAX: (415) 576-0300
      INFORMATION FOR SEQ ID NO: 2:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 776 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      US-08-082-849B-2

      Alignment Scores:
      Pred. NO.: 19.5      Length: 776
      Score: 48.50      Matches: 9
      Percent Similarity: 53.85%      Conservative: 5
      Best Local Similarity: 34.62%      Mismatches: 7
      Query Match: 29.57%      Indels: 5
      DB: 1      Gaps: 1

      US-09-513-151-3_COPY_1121_1210 (1-90) x US-08-082-849B-2 (1-776)

      QY 24 CATTGGGATGCGGANG-----GGCAGCGCACATAAATCAATCCCA 68
          |||||  ::  |||
      Db 280 H1STPserAspserLeuSerCargluGluGlyLeuLysIleuGlnIlepro 299
          ::|||  ::|||
      Db 300 IlegluProLysLysASP 305

      RESULT 13
      PCT-US94-01624-2
      Sequence 2, Application PC/TUS9401624
      GENERAL INFORMATION:
      APPLICANT: Leppla, Stephen H.
      APPLICANT: Klimpel, Kurt R.
      APPLICANT: Arora, Naveen
      APPLICANT: Singh, Yogendra
      APPLICANT: Nichols, Peter J.
      TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
      TITLE OF INVENTION: RELATED METHODS
      NUMBER OF SEQUENCES: 31
      CORRESPONDENCE ADDRESS:
      ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW
      STREET: Stewart Street Tower, 20th Floor, One Market
      STREET: Plaza
      CITY: San Francisco
      STATE: CA
      COUNTRY: USA
      ZIP: 94105

      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: PCT/US94/01624
      FILING DATE: June 25, 1993
      CLASSIFICATION:
      ATTORNEY/AGENT INFORMATION:
      NAME: Weber, Kenneth A.
      REGISTRATION NUMBER: 31,677
      REFERENCE/DOCKET NUMBER: 15280-115
      TELECOMMUNICATION INFORMATION:

```

TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 776 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-01624-2

Alignment Scores:

Pred. No.:	Length:
Score: 19.5	776
Percent Similarity: 48.50	Matches: 9
Best Local Similarity: 53.85%	Conservative: 5
Query Match: 34.62%	Mismatches: 7
	Indels: 5
DB: 5	Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x PCT-US94-01624-2 (1-776)

QY 24 CATTGGGATCGCGAATG-----GGCAGCGCACATTAATCCAAATCCCA 68

DB 280 HSTRPserAspserLeuSerclnuglYargGlyLeuLeuLysLysLennlilePro 299

QY 69 CTGAACCACTGAAGAA 86

DB 300 lIeGIuProlYlYsasp 305

RESULT 14

US-09-735-934A-2

Sequence 2, Application US/09735934A

Patent No. 6372468

GENERAL INFORMATION:

APPLICANT: LI, JiaYin et al

TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

FILE REFERENCE: CLO00851

CURRENT APPLICATION NUMBER: US/09/735,934A

CURRENT FILING DATE: 2000-12-14

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 878

TYPE: PRT

ORGANISM: Homo sapiens

US-09-735-934A-2

Alignment Scores:

Pred. No.:	Length:
Score: 24	878
Percent Similarity: 48.00	Matches: 8
Best Local Similarity: 70.59%	Conservative: 4
Query Match: 47.06%	Mismatches: 5
	Indels: 0
DB: 4	Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-735-934A-2 (1-878)

QY 66 GGATTGGATTATGTCGCTGCCCATTCGCGATCCCAATGATGATCG 16

DB 173 GlyLeuAsnTYrHISLysArgCysAlaPheSerileProAsnAsnCysSer 189

RESULT 15

US-09-135-782-4

Sequence 4, Application US/09135782

Patent No. 6027929

GENERAL INFORMATION:

APPLICANT: Xu, Shuang-Yong

TITLE OF INVENTION: Method For Cloning And Producing The NspI Restriction

TITLE OF INVENTION: Endonuclease In E. coli And Purification Of The

TITLE OF INVENTION: Recombinant NspI Restriction Endonuclease

FILE REFERENCE: NEB-143

CURRENT APPLICATION NUMBER: US/09/135,782

CURRENT FILING DATE: 1998-08-18
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 4
LENGTH: 244
TYPE: PRT
ORGANISM: No. 6027929loc sp.
US-09-135-782-4

Alignment Scores:

Pred. No.:	Length:
Score: 22.4	244
Percent Similarity: 47.50	Matches: 9
Best Local Similarity: 51.85%	Conservative: 5
Query Match: 33.33%	Mismatches: 12
	Indels: 1
DB: 3	Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-135-782-4 (1-244)

QY 1 TGTGACCTCTGTGATGATCATCATTTGGGATCGCGATGGCGACGACATAAATCC 60

DB 76 CysGIInThrCysGlnGlnInThrArgLeuGlyAspArg--TrpThrGlyGlnIleLysSer 94

QY 61 AATCCCACTTGACCAACTG 81

DB 95 AlaAsnSerAlaSerAsnIle 101

Search completed: April 21, 2003, 18:52:19

Job time : 2.97137 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 21, 2003, 18:47:50 ; Search time 4.20225 Seconds
(without alignments)
8825.857 Million cell updates/sec

Title: US-09-513-151-3_COPY_1121_1210

Perfect score: 164
Sequence: 1 TGTGACCTCTGTGATGCAAT.....TGAACCACTGAGAAAAAGA 90

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
	Ygapop 10.0 , Ygapext 0.5
	Rgapop 6.0 , Rgapext 7.0
	Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+2p.model -DEV=xih
-O=/cqr2_1/uspro.spool/US09513151/runat_15042003_141144_26380/app.query.fasta_1.2446
-DB=SPREMBL_21 -OFMT=fasta -SUFFIX=n2p.rspt -MINMATCH=0.1 -IOBCL=0
-LOOPEXT=0 -UNITS=DITS -START=1 -END=1 -MATRIX=Dlosum2 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=PCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=PLO -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09513151.qcgn.1.1.125.qunat.15042003_141144_26380 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGOQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	164	100.0	324	4	Q96FJ3		Q96FJ3 homo sapien

2	164	100.0	326	4	Q9NXT7	Q9NXT7 homo sapien
3	164	100.0	435	4	Q96LA5	Q96LA5 homo sapien
4	164	100.0	467	4	Q9H3H1	Q9H3H1 homo sapien
5	157	95.7	326	11	Q9D1H5	Q9D1H5 mus musculus
6	64	39.0	294	4	Q9UL40	Q9UL40 homo sapien
7	63	38.4	294	11	Q9R0B7	Q9R0B7 mus musculus
8	62	37.8	524	13	Q42147	Q42147 xenopus lae
9	61	37.2	314	5	Q9VHM6	Q9VHM6 drosophila
10	60	36.6	430	5	Q9GYG3	Q9GYG3 caenorhabdi
11	59	36.0	478	10	Q8W374	Q8W374 oryza sativ
12	58.5	35.7	239	10	Q8S9Z6	Q8S9Z6 oryza sativ
13	58	35.4	558	5	Q94540	Q94540 drosophila
14	58	35.4	578	5	Q9VA29	Q9VA29 drosophila
15	58	35.4	635	5	Q95026	Q95026 drosophila
16	58	35.4	884	5	Q9V0V6	Q9V0V6 drosophila
17	57.5	35.1	935	5	Q9VZ07	Q9VZ07 drosophila
18	57	35.2	128	12	Q8A660	Q8A660 simlan viru
19	57	34.8	818	4	Q8WU72	Q8WU72 homo sapien
20	57	34.8	837	4	Q8RTG3	Q8RTG3 homo sapien
21	57	34.8	842	4	Q9H865	Q9H865 homo sapien
22	57	34.8	902	13	Q8UWC5	Q8UWC5 gallus gall
23	57	34.8	1052	11	Q88532	Q88532 mus musculi
24	57	34.8	1057	4	Q96KR1	Q96KR1 homo sapien
25	56	34.1	346	5	Q9VHM4	Q9VHM4 drosophila
26	56	34.1	463	3	Q9S811	Q9S811 schizosacch
27	56	34.6	481	16	P72956	P72956 synechocyst
28	56	34.1	1053	4	Q95625	Q95625 homo sapien
29	56	34.1	1623	5	Q45019	Q45019 caenorhabdi
30	55.5	33.8	251	10	Q9AWK1	Q9AWK1 oryza sativ
31	55.5	33.8	336	5	Q8TOY0	Q8TOY0 apis cerana
32	55.5	33.8	539	5	Q9V4M9	Q9V4M9 drosophila
33	55.5	33.8	600	4	Q96MN7	Q96MN7 homo sapien
34	55	33.5	379	2	Q48738	Q48738 lactococcus
35	55	33.5	670	2	Q9AKS7	Q9AKS7 drosophila
36	55	34.0	652	5	Q9AEF6	Q9AEF6 leptospira
37	55	34.0	670	2	Q8VTY1	Q8VTY1 leptospira
38	55	34.0	670	2	Q8VTU9	Q8VTU9 leptospira
39	55	33.5	789	5	Q9VBX4	Q9VBX4 drosophila
40	55	33.5	3140	12	Q84925	Q84925 plum pox vi
41	54	32.9	199	4	Q96NC0	Q96NC0 homo sapien
42	54	32.9	199	11	Q9CEP7	Q9CEP7 mus musculi
43	54	32.9	845	11	Q8V8E2	Q8V8E2 mus musculi
44	53	32.3	225	16	Q8ZCD7	Q8ZCD7 yersinia pe
45	53	32.7	298	17	Q9H1L4	Q9H1L4 thermoplasm

ALIGNMENTS

RESULT 1

Q96FJ3 PRELIMINARY: PRT: 324 AA.

AC Q96FJ3
ID Q96FJ3
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Similar to crna isopenlentyipropiosate transferase.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;
RA Strausberg R.;
RT Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC010741; AAH0741.1;
DR InterPro: IPR002627; IPTT;
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF01715; IPTT; 1.
DR ProDom: PD004674; IPTT; 1.
DR SMART: SM00355; Znf_C2H2; 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KM Transferase.

SO SEQUENCE 324 AA: 37223 MW: 1E6835D7C09126A9 CRC64;

Alignment Scores:

Pred. No.: 7,31e-17 Length: 324
Score: 164.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x Q96FJ3 (1-324)

OY 1 TGTGACCTCTGTGATGCAATCATCTGGGATCGGCAATGGCAGCGCATAAATCC 60

DB 254 CysAspLeucCysAspArgIleIleIleGlyAspArgGluTrpAlaIleHisIleLysSer 273

OY 61 AAATCCCACTTGAAACCACTGAAGAAAGA 90

DB 274 LysSerHisLeuAsnGlnLeuLysArg 283

RESULT 2

O9NXT7 PRELIMINARY: PRT; 326 AA.

AC Q9NXT7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CDNA FLJ20061 f1s, clone COL01383.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Kawabata A., Hkiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shidhara T.,
RA Tanaka I., Nakamura Y., Isogai T., Sugano S.;
RT "NEO human cDNA sequencing project."
RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF000068; BAA90923.1;
DR InterPro: IPR002627; IPT: -
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF01715; IPT: 1.
DR ProDom: PD004674; IPT: 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
SQ SEQUENCE 326 AA: 37435 MW: EA83F0F664B7ACE CRC64;

Alignment Scores:

Pred. No.: 7,31e-17 Length: 326
Score: 164.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x Q9NXT7 (1-326)

OY 1 TGTGACCTCTGTGATGCAATCATCTGGGATCGGCAATGGCAGCGCATAAATCC 60

DB 256 CysAspLeucCysAspArgIleIleIleGlyAspArgGluTrpAlaIleHisIleLysSer 275

OY 61 AAATCCCACTTGAAACCACTGAAGAAAGA 90

DB 276 LysSerHisLeuAsnGlnLeuLysArg 285

RESULT 3

O96L45 PRELIMINARY: PRT; 435 AA.

AC Q96L45;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE tRNA isopentenyl transferase (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxId=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE-21444833; PubMed-11560893;

RA Lemieux J., Lakowski B., Webb A., Meng Y., Ubach A., Bussiere F.,

RA Barnes T., Hekimi S.;

RT "Regulation of Physiological Rates in Caenorhabditis elegans by a

RT tRNA-Modifying Enzyme in the Mitochondria.";

RL Genetics 159:147-157(2001).

DR EMBL; AF052768; AAL14107.1; -

DR InterPro: IPR002627; IPT: -

DR Pfam: PF01715; IPT: 1.

DR ProDom: PD004674; IPT: 1.

DR SMART; SM00355; Znf_C2H2_1.

DR TIGRFAMs; TIGR00174; mlaa; 1.

DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.

KW transferase.

FT NON_TER

SQ SEQUENCE 435 AA: 48948 MW: 2279AE7C2D999FF1 CRC64;

Alignment Scores:

Pred. No.: 7,4e-17 Length: 435
Score: 164.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x Q96L45 (1-435)

OY 1 TGTGACCTCTGTGATGCAATCATCTGGGATCGGCAATGGCAGCGCATAAATCC 60

DB 365 CysAspLeucCysAspArgIleIleIleGlyAspArgGluTrpAlaIleHisIleLysSer 384

OY 61 AAATCCCACTTGAAACCACTGAAGAAAGA 90

DB 385 LysSerHisLeuAsnGlnLeuLysArg 394

RESULT 4

O9H3H1 PRELIMINARY: PRT; 467 AA.

AC Q9H3H1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE tRNA isopentenylpyrophosphate transferase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.

RA MEDLINE-20564178; PubMed-11111046;

RA Golovko A., Hjalms G., Sitbon F., Nicander B.;

RT "Cloning of a human tRNA isopentenyl transferase.";

RL Gene 238:85-93(2000).

DR EMBL; AF074918; AAG31324.1; -

DR InterPro: IPR002627; IPT: -

DR InterPro: IPR000822; Znf_C2H2.

DR Pfam: PF01715; IPT: 1.

DR ProDom: PD004674; IPT: 1.

DR TIGRFAMs; TIGR00174; mlaa; 1.

DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.

KW transferase.

SQ SEQUENCE 467 AA: 52725 MW: 634469919D7F56A5 CRC64;

Alignment Scores:

Pred. No.: 7,42e-17 Length: 467


```

DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE Double-stranded RNA-binding zinc finger protein JAZ.
GN ZFP346.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99419013; PubMed=10488071;
RA Yang M., May W.S., Ito T.;
RT "JAZ requires the double-stranded RNA-binding zinc finger motifs for
RT nuclear localization."
RL J. Biol. Chem. 274:27399-27406(1999).
DR EMBL: AF083339; AAD52017.1;
DR MGP: MGI:1349417; Zfp346.
DR InterPro: IPR000822; Znf_C2H2.
DR InterPro: IPR003604; Znf_U1.
DR Pfam: PF00096; Zf-C2H2; 3.
DR SMART: SM00355; Znf_C2H2; 4.
DR SMART: SM00451; Znf_U1; 4.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN.4.
KW DNA-binding; Zinc-finger.
SO SEQUENCE 294 AA; 32698 MW; 6BD916262EDBA71E CRC64;

Alignment Scores:
Pred. No.: 0.57 Length: 294
Score: 63.00 Matches: 10
Percent Similarity: 65.52% Conservative: 9
Best Local Similarity: 34.48% Mismatches: 10
Query Match: 38.41% Indels: 0
DB: 11 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x Q9R0B7 (1-294)
QY 1 TGTGACCTGTGTGATCATCATTCATGCGGATCGGACGACGCATATAATCC 60
DB 75 CysLysValcysCysAlaMetLeuIleSerGlnLysLeuAlaHisTyrGlnSer 94
QY 61 AAATCCCACTTGACCACTGAAGAAA 87
DB 95 LysLysHisAlaAsnLysValLysArg 103

RESULT 8
042147 PRELIMINARY: PRT; 524 AA.
AC 042147;
DT 01-JAN-1998 (TRENBLREL. 05, Created)
DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)
DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE DSRBP-2Fa.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97415754; PubMed=9268652;
RA Finerty P.J., Jr., Bass B.L.;
RT "A Xenopus zinc finger protein that specifically binds dsRNA and RNA-
RT DNA hybrids."
RL J. Mol. Biol. 271:195-208(1997).
DR EMBL: AF005083; AAC60260.1;
DR InterPro: IPR000822; Znf_C2H2.
DR InterPro: IPR000690; Znf_Matrin.
DR InterPro: IPR003604; Znf_U1.
DR Pfam: PF00096; Zf-C2H2; 5.
DR SMART: SM00355; Znf_C2H2; 7.
DR SMART: SM00451; Znf_U1; 7.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 7.
KW DNA-binding; Metal-binding; Zinc-finger.
SO SEQUENCE 524 AA; 55588 MW; 104D86038ADB8D57 CRC64;

```

```

Alignment Scores:
Pred. No.: 0.84 Length: 524
Score: 62.00 Matches: 8
Percent Similarity: 65.52% Conservative: 11
Best Local Similarity: 27.59% Mismatches: 10
Query Match: 37.80% Indels: 0
DB: 13 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x 042147 (1-524)
QY 1 TGTGACCTGTGTGATCATCATTCATGCGGATCGGACGACGCATATAATCC 60
DB 36 CysLysValcysSerAlaValLeuIleSerGlnLysLeuAlaHisTyrGlnSer 55
QY 61 AAATCCCACTTGACCACTGAAGAAA 87
DB 56 ArgLysHisAlaAsnLysValLysArg 64

RESULT 9
Q9VHM6 PRELIMINARY: PRT; 314 AA.
AC Q9VHM6;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE CG11762 protein.
GN CG11762.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.T., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtils K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheel F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svyrskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

```

RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003680; AAF54277.1; -.
 DR HSSP: P15822; IBO.
 DR FlyBase: FBgn0037618; CG11762.
 DR InterPro: IPR000822; znf.C2H2.
 DR InterPro: IPR000690; znf.mattin.
 DR Pfam: PF00096; znf.C2H2; 5.
 DR SMART: SM00355; znf.C2H2; 5.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 4.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 5.
 DR DNA-binding; Metal-binding; Zinc-finger.
 SO SEQUENCE 314 AA; 36509 MW; C783E29C050F5D57 CRC64;

Alignment Scores:
 Pred. No.: 1.18 Length: 314
 Score: 61.00 Matches: 11
 Percent Similarity: 50.00% Conservative: 4
 Best Local Similarity: 36.67% Mismatches: 15
 Query Match: 37.20% Indels: 0
 DB: Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x Q9VHM6 (1-314)

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCAGCGACATATAATCC 60
 DB 283 CysaspilicysasparsSerPheGlnArgLysAlaHisLeuValThrHisThrArgSer 302

OY 61 AAATCCCACTGAACCACTGAGCAAGAAAGA 90
 DB 303 MetethisthLeuGlnAsnValLysLysGln 312

RESULT 10
 O9GYG3 PRELIMINARY; PRT; 430 AA.
 AC O9GYG3: 095UF6;
 DT 01-MAR-2002 (TREMBLrel. 16, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 49.5 kDa protein (TRNA isopentenyl transferase).
 OS ZC395.6 OR GRO-1.
 GN *Caenorhabditis elegans*.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode *C. elegans*: a platform for
 RT investigating biology. The *C. elegans* Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RC SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Connell M.;
 RT "The sequence of *C. elegans* cosmid ZC395.";
 RL Submitted (NOV-1994) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RC SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RC SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA MEDLINE=21444833; PubMed=11560893;
 RA Lemieux J., Lakowski B., Webb A., Meng Y., Ubach A., Bussiere F.,
 RA Barnes T., Hekimi S.;
 RT "Regulation of Physiological Rates in *Caenorhabditis elegans* by a
 RT tRNA-Modifying Enzyme in the Mitochondria.";

RL Genetics 159:147-157(2001).
 DR EMBL: U13642; AAG00042.2; -.
 DR DR AY052273; AAL14112.1; -.
 DR InterPro: IPR002627; IPT.
 DR InterPro: IPR000822; znf.C2H2.
 DR Pfam: PF01715; IPT; 1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 DR Hypothetical protein; DNA-binding; Transferase; Zinc-finger.
 SO SEQUENCE 430 AA; 49548 MW; 107E95095C81A2B9 CRC64;

Alignment Scores:
 Pred. No.: 1.72 Length: 430
 Score: 60.00 Matches: 9
 Percent Similarity: 53.33% Conservative: 7
 Best Local Similarity: 30.00% Mismatches: 14
 Query Match: 36.59% Indels: 0
 DB: Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x Q9VY3 (1-430)

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCAGCGACATATAATCC 60
 DB 394 CysglnilicysasnileSerMetThrGlyLysAspAsnTrpGlnLysHisilespely 413

OY 61 AAATCCCACTGAACCACTGAGCAAGAAAGA 90
 DB 414 LysLysHisLysHisHisAlaLysGlnLys 423

RESULT 11
 O8W374 PRELIMINARY; PRT; 278 AA.
 AC O8W374:
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 29.2 kDa protein.
 GN OSJNBA0029C15.17.
 OS *Oryza sativa* (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
 RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
 RA Tsirlin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pal G.,
 RA Vanden S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
 RA Salzberg S.L., White O., Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSJNBA0029C15 genomic sequence.";
 RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC087182; AAL59020.1; -.
 DR InterPro: IPR000822; znf.C2H2.
 DR InterPro: IPR003604; znf.U1.
 DR SMART: SM00451; znf.U1; 2.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_2.
 DR Hypothetical protein.
 SO SEQUENCE 278 AA; 29241 MW; 7674AC0C3341DAC CRC64;

Alignment Scores:
 Pred. No.: 2.42 Length: 278
 Score: 59.00 Matches: 10
 Percent Similarity: 57.14% Conservative: 6
 Best Local Similarity: 35.71% Mismatches: 12
 Query Match: 35.98% Indels: 0
 DB: Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x O8W374 (1-278)

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCAGCGACATATAATCC 60
 DB 230 CysglnleucycysaspvalleuAlaSerGlnleuAsnValAlaIleHisThrLagly 249

```

QY 61 AATCCCACTTGACCACTGAG 84
Db 250 LysGlnHisLeuHisArgLeuArg 257

RESULT 12
Q8S926 PRELIMINARY: PRT: 439 AA.
ID Q8S926;
AC Q8S926;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE Similar to tRNA Isopenentenyltransferase.
GN OJ1656.A11.14.
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa niponbare(GA3) genomic DNA, chromosome 1, BAC
clone:OJ1656.A11."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, AP003448; BAB5325.1; -.
KW Transferase.
SO SEQUENCE 439 AA; 49646 MW; 74BE689EAA2EEF95 CRC64;

Alignment Scores:
Pred. No.: 2.96 Length: 439
Score: 58.50 Matches: 9
Percent Similarity: 60.00% Conservative: 9
Best Local Similarity: 30.00% Mismatches: 11
Query Match: 35.67% Indels: 1
DB: 10 Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x Q8S926 (1-439)
QY 1 TGTGACCTCTGTGAT---CGAATCATTCGCGAATGGCGACGACATAA 57
Db 399 CysGlnHisArgValLeuArgGlyThrHisGluTrpGlnHisLysGln 418
QY 58 TCCAATCCCACTTGACCACTGAG 87
Db 419 GlyArgCysHisArgLysArgValGlnArg 428

RESULT 13
Q94540 PRELIMINARY: PRT: 558 AA.
ID Q94540;
AC Q94540;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DE 1-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Zinc-finger protein ZNF22D.
GN ZNF22D OR CG5215.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97134673; PubMed=8980233;
RA Sauer F., Massaman D.A., Rubin G.M., Tjian R.;
RT "AFR115 mediate activation of transcription in the Drosophila
embryo."
RL Cell 87:1271-1284(1996);
DR EMBL, U73125; AAB40721.1; -.
DR FlyBase; FBgn0017453; ZNF22D.
DR InterPro; IPR000822; Znf_C2H2.
DR InterPro; IPR000690; Znf_matin.

```

```

DR InterPro; IPR003604; Znf_U1.
DR Pfam; PF00096; zf-C2H2; 3.
DR SMART; SM00355; ZNF_C2H2; 3.
DR SMART; SM00451; ZNF_U1; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.
KW DNA-binding; Zinc-finger.
SO SEQUENCE 558 AA; 61157 MW; 5DC76C1BEF2DC784 CRC64;

Alignment Scores:
Pred. No.: 3.59 Length: 558
Score: 58.00 Matches: 10
Percent Similarity: 53.57% Conservative: 5
Best Local Similarity: 35.71% Mismatches: 13
Query Match: 35.37% Indels: 0
DB: 5 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x Q94540 (1-558)
QY 1 TGTGACCTCTGTGATCGAATCATTCGCGAATGGCGACGACATAAATCC 60
Db 178 CysGlnLeuCysAspValThrCysThrGlyThrAspAlaTyrAlaAlaHisValArgGly 197
QY 61 AATCCCACTTGACCACTGAG 84
Db 198 AlaLysHisGlnLysValValLys 205

RESULT 14
Q9VA29 PRELIMINARY: PRT: 578 AA.
ID Q9VA29;
AC Q9VA29;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE CG12071 protein.
GN CG12071.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bales R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butlis J.K., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glöcker A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ijzerman C.,
RA Jalaal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pauley J.M.,
RA Palazolo M., Peltan G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeffer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

```

RA Splter E, Spreading A.C., Stapleton M., Strong R., Sun E.
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zavadil J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Ra Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu S.,
Ra Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003775; AAF57095.1; -.
DR HSSP; P07248; 2ADR.
DR FlyBase; FBgn0039808; CG12071.
DR InterPro; IPR000823; Znf-C2H2.
DR InterPro; IPR000966; zf-C2H2_3.
DR Pfam; PF000003; Znf-C2H2_1.
DR ProDom; PD000003; Znf-C2H2_1.
DR SMART; SM00355; Znf-C2H2_3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_3.
KW DNA-binding; Metal-binding; Zinc-finger.
SO SEQUENCE 578 AA; 64636 MW; 19966AD07BFA43FE CRC64;

Alignment Scores:

Pred. No.:	3.59	Length:	57
Score:	58.00	Matches:	11
Percent Similarity:	62.50%	Conservative:	4
Best Local Similarity:	45.83%	Mismatches:	7
Query Match:	35.37%	Indels:	2
DB:	5	Gaps:	1

US-09-513-151-3_COPY_1121_1210 (1-90) x Q9VA29 (1-578)

```

0Y      1  TGTGACCCCTGTGATGCATCATCATTTGGGATGCGCATGTGGCAGCGACATATAAATTC 60
        |||:::|||||:::|||||:::|||:::|||||:::|
Db      315  Cysglutenucysglarygmecphesetsernarygsplutrpserllehs-----Ala 333
        |||:::|||||:::|||||:::|||:::|||||:::|

0Y      61  AAATCCCACTTG 72
        |||:::|||||:::|||||:::|||:::|||||:::|
Db      333  LysSerHisIeu 336
        |||:::|||||:::|||||:::|||:::|||||:::|

```

RESULT 15

ID	Q95U26	PRELIMINARY; PRT; 635 AA.
AC	Q95U26;	
DT	01-DEC-2001	(TREMBLrel. 19, Created)
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)

05 *Drosophila melanogaster* (Fruit fly).
06 *Eukaryota*, Metazoa, Arthropoda, Tracheata, Hexapoda, Insecta;
07 *Pterygota*, Neoptera, Endopterygota, Diptera, Brachycera, Muscomorpha
08 *Ephyridroidea*, *Drosophilidae*, *Drosophila*.
09 NCBI_taxid=7227;

RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY.
 RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuccio J., Paolel J., Paragas V., Park S., Phouaneavong S., Wan K.,
 RA Yu C., Lewis S.E., Rulins G.M., Celisner S.;
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY058357; AAL13586.1; -;
 DR FlyBase: FBgn0017453; Znf2D.
 DR InterPro: IPR000822; Znf_C2H2.
 DR Pfam: PF00096; zf_C2H2_3.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.
 KW DNA-binding; Zinc-finger.
 SO SEQUENCE 635 AA; 68457 MW; 097515874204ECDE CRC64;

Alignment Scores:

Pred. No.:	3.61	Length:	635
Score:	58.00	Matches:	10

```
Percent Similarity: 53.57% Conservative: 5
Best Local Similarity: 35.71% Mismatches: 1
Query Match: 35.37% Indels: 0
DB: 5 Gaps: 0
```

US-09-513-151-3 COPY_1121_1210 (1-90) x Q95U26 (1-635)

qy 1 TGTGACCTCTGTGATCGAATCATCTATGGGGATCCGGAATGGGACGGCAGACATAAAATCC 60
|||...||| |||...|||

Db 255 CysGluLeuCysAspValThrCysThrGlyThrAspAlaTyrAlaAlaHisValArgGly 274

61 AAATCCCACTTGAACCAACTGAAG 84

Db 275 ALaLySHiGlnLyVaIValLyS 282

Search completed: April 21, 2003, 18:55:43
Job time : 7.20225 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 18:21:20 ; Search time 2.11168 Seconds
(without alignments)
8194.498 Million cell updates/sec

```
Title: US-09-513-151-3_COPY_1121_1210
Perfect score: 164
Sequence: 1 TGTGACCTCTGTGATCGAAT.....TGACCAACTGAGAAACA 90
```

Scoring table:		BloSUM62
Xgapop	10.0 ,	Xgapext 0.5
Ygapop	10.0 ,	Ygapext 0.5
Fgapop	6.0 ,	Fgapext 7.0
Delop	6.0 ,	Delext 7.0

```
Searched:      283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 566448
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Maximum Match 100%
Listing first 45 summaries

```

Command line parameters:
-MODEL=frimatt_n2p_model -DEV=xlh
-Q=-cgn2-11/USPFO.spool/US09513151/runat.15042003.14114.26390/app.query.fasta-1.2446
-DB=erit-73 -QFMT=fastan -SUFFIX=n2p.rpt -MINMATCH=0-1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blissom62 -TRANS=align40.csl -LIST=45
-DLOCAL=100 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPUT=pct -NORM=act -HEAPSIZE=500 -MTEXT=2000000000
-USER=US09513151 -CGCN.1.1.58.0runat.15042003.14114.26390 -NCPY=6 -ICPU=3
-NO.XLPHYT -NO.MAP -LARGEOBJECT -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THRAD=1 -XGADP=10 -XGAPEXT=0.5 -FGADP=6 -FGAPEXT=7
-XGADP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

```
Database :
1: plr1:*
2: plr2:*
3: plr3:*
4: plr4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	36.6	433	2	T27538	hypothetical protein
2	59	36.0	428	2	S67176	tRNA-isopentenyltr
3	58.5	35.7	453	2	B69504	seryl-tRNA synthet
4	58	35.4	301	2	S56211	probable membrane
5	57	35.2	180	2	I48129	Xel9 (escapes X-1
6	57	35.2	180	2	I84689	escapes X-inactiva
7	57	34.8	1052	2	T14343	zinc finger RNA bi
8	57	35.2	1560	2	I54361	SMCX protein - hum
9	56	34.1	263	2	G87721	protein ccl123.3 (l
10	56	34.1	463	2	T41390	zinc finger protei
11	56	34.6	481	2	S74834	plp-1 protein - S
12	55	33.5	379	2	C57505	integrase - lactoc
13	55	33.5	3140	2	S47508	genome polypeptid
14	54	32.9	222	2	H64422	type II restrictio

45	44	43	42	41	40	39	38	37	36	35	34	33	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14
29.9	29.9	29.9	29.9	29.9	30.2	30.6	30.6	30.2	30.2	30.2	30.2	30.5	30.5	30.9	30.5	30.5	30.5	30.5	30.5	30.5	30.5	31.1	31.1	31.5	31.4	31.8	31.7	32.3	32.6	32.5	32.6
395	313	157	61	1817	479	340	340	214	172	113	364	574	200245	548	514	386	367	260	257	108	104	407	389	325	1044	63	1	433	225	232	
154232	664433	296833	T41856	H71611	T05588	T05964	T05964	H72377	T44356	H84166	AB0835	T00245	C64438	C64438	C44154	AE3582	AE3582	B45061	B45061	D96543	J74901	S40990	S40990	CY0282	DVBEX3	E37052	GNV34	AE0371	H84028		

DNA repair protein
 probable carboxypeptidase
 steroid hormone receptor
 genome polyprotein
 hypothetical protein
 transatlational elongation factor
 hypothetical protein
 hypothetical protein
 ZK686.4 protein - predicted
 phosphoprotein phosphatase
 arsenical resistant protein
 unknown protein (110 kDa)
 granzyme A (EC 3.4.21.1)
 granzyme A (EC 3.4.21.1)
 carboxine oxidase
 xylase operon transactivator
 hypothetical protein
 phenylalanine tRNA synthetase
 reverse transcriptase
 large repetitive protein
 hypothetical protein
 probable dCpF deaminase
 transcription regulator
 probable iron uptake protein
 cellulase (EC 3.2.2.6)
 cathepsin B mRNA 3' UTR
 ACNPNV orf122 - Borna disease virus
 hypothetical protein
 hypothetical protein
 C2H2 zinc finger protein

ALIGNMENTS

RESULT 1

hypothetical protein ZC395.6 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*

C:\Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:\Accession: T27538

submitted to the EMBL Data Library, August 1994

A; Reference number:

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Residues: 1-433 <

A; Experimental source: strain Bristol N2; clone ZC395

A; Gene: CESP:ZC395.6

A; Introns: 43/1; 201/1; 266/3; 310/2; 341/3; 377/3; 410/2

Alignment Scores:

Pred. No.:	1.21	Length:	43
Score:	60.00	Matches:	9
Percent Similarity:	53.33%	Conservative:	7
Best Local Similarity:	30.00%	Mismatches:	14
Query Match:	36.59%	Indels:	0
DB:	2	Gaps:	0

US-09-513-151-3_COPY_1121_1210 (1-90) x T27538 (1-433)

1 TGTGACCTCTGTGATCGAATCATTCATTGGGGATCGCGAATGGGCAGCGCACATAAAATCC 60

Db 397 Cysgluilecysasnilesermetthr glylysaspasnttrpglnlyshsileasp gly 410

61 AAATCCACATTGAACCAACTGAAAGAAAGA 90

Db 417 LysLysHisHisAlaLysGlnLys 426

RESULT 2

S67176

tRNA isopentenyltransferase (EC 2.5.1.8) - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: protein O547w; protein YOR274w

C:Species: *Saccharomyces cerevisiae*

C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 21-Jul-2000

C:Accession: S67176; A26717; S72045

R:Cheret, G.; Sor, F.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67169

A:Accession: S67176

A:Molecule type: DNA

A:Residues: 1-428 <CH>

A:Cross-references: EMBL:Z75182; NID:g1420613; PID:e252418; PID:g1420614; MIPS:YOR274w

A:Experimental source: strain S288C

R:Najarian, D.; Dhanich, M.E.; Martin, N.C.; Hopper, A.K.

Mol. Cell. Biol. 7, 185-191, 1987

A:Title: DNA sequence and transcript mapping of MOD5: features of the 5' region which su

A:Reference number: A26717; MUID:87172703; PMID:3031457

A:Accession: S72045

A:Molecule type: DNA

A:Residues: 1-374, R', 376-428 <NA>

A:Cross-references: EMBL:M15991

R:Cheret, G.; Bernardi, A.; Sor, F.

Yeast 12, 1059-1064, 1996

A:Title: DNA sequence analysis of the VP1-SNF2 region on chromosome XV of *Saccharomyces*

A:Reference number: S72039; MUID:97051594; PMID:8896271

A:Accession: S72045

A:Molecule type: DNA

A:Residues: 1-428 <CH>

A:Cross-references: EMBL:X89633; NID:g1279694; PIDN:CA61780.1; PID:g1419759

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995

C:Genetics:

A:Gene: SGD:MOD5

A:Cross-references: MIPS:YOR274w; SGD:S0005800

A:Map position: 15R

C:Keywords: transferase

Alignment Scores:

Pred. No.:	1 72	Length:	428
Score:	59.00	Matches:	10
Percent Similarity:	57.14%	Conservative:	10
Best Local Similarity:	28.57%	Mismatches:	9
Query Match:	35.98%	Indels:	6
DB:	2	Gaps:	1

US-09-513-151-3_COPY_1121_1210 (1-90) x S67176 (1-428)

QY 1 TGTGACCTCTGT-----GATGCAATCATCTATGGGATCGCAATGG 42

Db 375 CYSASNValCysrgrsnaAlaaspGlyLysASNValAlaAlaLeuGlyLysuYrTTP 394

QY 43 GCAAGCCACATTAATCCAAATCCACTGACCAAGCAAGAA 87

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

Db 395 LysIleHisLeuGlySerArgHisLysSerASNLeuLysArg 409

```

A:Accession: I48129
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-180 <RES>
A:Cross-references: GB:I29564; NID:9535460; PIDN:AAA62383.1; PID:9535461
C:Genetics:
A:Gene: Xel69
C:Superfamily: human retinoblastoma binding protein 2

Alignment Scores:
Pred. No.: 3 37 Length: 180
Score: 57.00 Matches: 9
Percent Similarity: 64.00% Conservative: 7
Best Local Similarity: 36.00% Mismatches: 3
Query Match: 35.19% Indels: 6
DB: 2 Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x I48129 (1-180)

QY 80 AGTGGTTCAGTGGATTGATTTATGTCGCGCCCATTCGCGATCCCAATGATG 21
      ::::::::::::::::::::|
Db 29 Alatrptpqlutrpasphtlryspheleucys-----Prolaucys 42
      ::::::::::::::::::::|

QY 20 ATTCGATCAGAGG 6
      ::::::::::::::|
Db 43 MetargserArgarg 47

RESULT 6
I84689
escapes X-inactivation - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 26-May-2000
C:Accession: I84689
R:Wu, J.; Salido, E.C.; Yen, P.H.; Mohandas, T.K.; Heng, H.H.; Tsui, L.C.; Park, J.; Cha
Nature Genet. 7, 491-496, 1994
A:Title: The murine Xel69 gene escapes X-inactivation like its human homologue.
A:Reference number: I48129; MUID:95038834; PMID:7951318
A:Accession: I84689
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-180 <RES>
A:Cross-references: GB:I29563; NID:9535683; PIDN:AAA62384.1; PID:9535684
C:Genetics:
A:Gene: Xel69
C:Superfamily: human retinoblastoma binding protein 2

Alignment Scores:
Pred. No.: 3 37 Length: 180
Score: 57.00 Matches: 9
Percent Similarity: 64.00% Conservative: 7
Best Local Similarity: 36.00% Mismatches: 3
Query Match: 35.19% Indels: 6
DB: 2 Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x I84689 (1-180)

QY 80 AGTGGTTCAGTGGATTGATTTATGTCGCGCCCATTCGCGATCCCAATGATG 21
      ::::::::::::::::::::|
Db 29 Alatrptpqlutrpasphtlryspheleucys-----Prolaucys 42
      ::::::::::::::::::::|

QY 20 ATTCGATCAGAGG 6
      ::::::::::::::|
Db 43 MetargserArgarg 47

RESULT 7
T14343
zinc finger RNA binding protein, chromosome-associated - mouse
C:Species: Mus musculus (house mouse)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14343
R:Headner, M.J.; Schumacher, J.M.; Lee, K.; Holdcraft, R.W.; Edelhoff, S.; Distche, C.;
Gene 228, 197-211, 1999
A:Title: Identification of ZFR, an ancient and highly conserved murine chromosome-assoc

```

```

A:Reference number: Z17994; MUID:99173884; PMID:10072773
A:Accession: T14343
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1052 <MEN>
A:Cross-references: EMBL:AF071059; NID:93293536; PID:93293537; PIDN:AAC25762.1
A:Experimental source: testis
C:Genetics:
A:Gene: Zfr
A:Map position: 15A
C:Keywords: zinc finger

Alignment Scores:
Pred. No.: 3 63 Length: 1052
Score: 57.00 Matches: 11
Percent Similarity: 53.57% Conservative: 4
Best Local Similarity: 39.29% Mismatches: 13
Query Match: 34.76% Indels: 0
DB: 2 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x T14343 (1-1052)

QY 1 TGTGACCTCTGTGATGCATCATTCGGGATCGCGATGGCAGCAGCATMAATCC 60
      ::::::::::::::|
Db 362 Cysgluleucysaspalsercytrhglyalaaspalatyralahistileargly 381
      ::::::::::::::|

QY 61 AATCCCACTTGAAACCACTGAG 84
      ::|
Db 382 Alalysghisglulysvalvallys 389

RESULT 8
I54361
SMCX protein - human
N:Alternate names: escapes X-chromosome inactivation
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 17-Mar-2000
C:Accession: I54361; I38235
R:Wu, J.; Ellison, J.; Salido, E.; Yen, P.; Mohandas, T.; Shapiro, L.J.
Hum. Mol. Genet. 3, 153-160, 1994
A:Title: Isolation and characterization of XE169, a novel human gene that escapes
A:Reference number: I54361; MUID:94214434; PMID:8162017
A:Accession: I54361
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1560 <MU>
A:Cross-references: GB:I25270; NID:9457136; PID:9457137
R:Agulnik, A.I.; Mitchell, M.J.; Mattei, M.G.; Borsani, G.; Avner, P.A.; Lerner, J
Hum. Mol. Genet. 3, 879-884, 1994
A:Title: A novel X gene with a widely transcribed Y-linked homologue escapes X-inac
A:Reference number: I38235; MUID:95038739; PMID:7951230
A:Accession: I38235
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 280-341, 'Y', 343-344 <AGU>
A:Cross-references: EMBL:229650; NID:9472994; PIDN:CAA82758.1; PID:9472995
C:Genetics:
A:Gene: GDB:SMCX; XE169; DXS1272E
A:Map position: Xp11.22-Xp11.21
C:Superfamily: human retinoblastoma binding protein 2

Alignment Scores:
Pred. No.: 3 7 Length: 1560
Score: 57.00 Matches: 9
Percent Similarity: 64.00% Conservative: 7
Best Local Similarity: 36.00% Mismatches: 3
Query Match: 35.19% Indels: 6
DB: 2 Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x I54361 (1-1560)

QY 80 AGTGGTTCAGTGGATTGATTTATGTCGCGCCCATTCGCGATCCCAATGATG 21
      ::::::::::::::|

```

Db 1234 AlATrPrpLutRPaSprHrLysPheLeuCys-----ProLeuCys 1247

OY 20 ATTCGATCAGAGG 6

Db 1248 MetArGerArArGaR 1252

RESULT 9

protein zc123.3 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C:Accession: G87721

R:Anonymous, The C. elegans Sequencing Consortium.

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.elegans/

A:Note: published errata appeared in Science 283, 35, 1999; Science 285, 2103, 1999; and

A:Accession: G87721

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-263 <STO>

A:Cross-references: GB:chr.I; PIDN:AB97603.1; PID:g2804499; GSPDB:GN00019; CESP:zc123.3

A:Note: contains similarity to C2H2-type zinc fingers

C:Genetics:

A:Gene: zc123.3

A:Map position: 1

Alignment Scores:

Pred. No.: 4.88 Length: 263

Score: 56.00 Matches: 9

Percent Similarity: 50.00% Conservative: 6

Best Local Similarity: 30.00% Mismatches: 15

Query Match: 34.15% Indels: 0

DB: 2 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x G87721 (1-263)

OY 1 TGTGACCTCTGTGATTCATCATTCGCGATGCGCATGCGCATATAATCC 60

Db 204 CysGlnuetysserGluThrPheAlaThrLysGlnAlaPheLeuSerHisLeuAsnSer 223

OY 61 AATATCCACTTGAACCACTGAGAAAGA 90

Db 224 AlaArgHisLeuGlnGlnAlaLysLysGln 233

RESULT 10

T41390

zinc finger protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000

C:Accession: T41390

R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, May 1998

A:Reference number: Z21990

A:Accession: T41390

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1463 <NR>

A:Cross-references: EMBL:AL023592; PIDN:CA19119.1; GSPDB:GN00068; SPDB:SPCC550.15c

A:Experimental source: strain 972h-; cosmid c550

C:Genetics:

A:Gene: SPDB:SPCC550.15c

A:Map position: 3

C:Superfamily: zinc finger protein SPCC550.15c

Alignment Scores:

Pred. No.: 5 Length: 463

Score: 56.00 Matches: 8

Percent Similarity: 62.07% Conservative: 10

Best Local Similarity: 27.59% Mismatches: 11

Query Match: 34.15% Indels: 0

DB: 2 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x T41390 (1-463)

OY 1 TGTGACCTCTGTGATTCATCATTCGCGATGCGCATGCGCATATAATCC 60

Db 72 CysGlnuValCysAsnLysPheTyrSerClnuLysAlaTyrSerHisMetAsnSer 91

OY 61 AATATCCACTTGAACCACTGAGAAAGA 87

Db 92 LysLysHisArgAspAsnLeuSerLys 100

RESULT 11

574934

pleb-1 protein - Synechocystis sp. (strain PCC 6803)

N:Alternate names: protein slr0687

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999

C:Accession: S74934

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Ye

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

A:Reference number: S74934; MUID:97061201; PMID:8905231

A:Accession: S74934

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1481 <KAN>

A:Cross-references: EMBL:D90902; GB:AB01339; NID:91652027; PIDN:BA16974.1; PID:dic

C:Genetics:

A:Gene: pleb-1

A:Start codon: GTG

C:Superfamily: response regulator homology

C:Keywords: phosphoprotein

F:10-122/Domain: response regulator homology <RRH>

F:61/Binding site: phosphate (Asp) (covalent) #status predicted

Alignment Scores:

Pred. No.: 5.01 Length: 481

Score: 56.00 Matches: 10

Percent Similarity: 61.54% Conservative: 6

Best Local Similarity: 38.46% Mismatches: 8

Query Match: 34.57% Indels: 2

DB: 2 Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x S74934 (1-481)

OY 80 AGTGTTCAGATGGAATTGATTTATGTCGCGCATGCGCATATCCCATGATG 21

Db 281 SerTrpLeuProTrpGluIleGluPheLeuCys-----HisLeuSerSerGlnValAla 298

OY 20 ATTCGATCAGAGGCTCA 3

Db 299 IleAlaIleGlnGlnSer 304

RESULT 12

C55205

Integrase - Lactococcus lactis transposon Tn5276

C:Species: Lactococcus lactis

C:Date: 05-May-1995 #sequence_revision 05-May-1995 #text_change 15-Oct-1999

C:Accession: C55205

R:Rauch, P.J.G.; de Vos, W.M.

J. Bacteriol. 176, 2165-2171, 1994

A:Title: Identification and characterization of genes involved in excision of the Ia

A:Reference number: A55205; MUID:94209213; PMID:8157585

A:Accession: C55205

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-379 <RAU>

A:Cross-references: GB:U27649; NID:94977770; PIDN:AAA72427.1; PID:9497773

C:Genetics:

Db 208 GluLeuAspHisValIleIleGlyAspArgThrPheValSer---LeuLysGluLys 226
Oy 64 TCCCACTG 72
|||
Db 227 GlyHisLeu 229

Search completed: April 21, 2003, 18:51:12
Job time : 5.11168 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 17:14:14 ; Search time 1.07696 Seconds
(without alignments)
6932.240 Million cell updates/sec

Title: US-09-513-151-3_COPY_1121_1210
Sequence: 1 TGTGACCTCTGTGATCGAAT.....TGACCACTGACAGAAAAGA 90

Scoring table:
BIOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_{n2p}.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09513151/unat_15042003_141143_26366/app_query.fasta.1.2446
-DB=SwissProt.40 -QFMT=fastan -SUPFIX=n2p.rsp -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09513151 -ECGN_1_1_25=etunal_15042003_141143_26366 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAR -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt.40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	63	38.4	845 1	MAT3_RAT
2	63	38.4	847 1	MAT3_HUMAN
3	59	36.0	428 1	MOD5_YEAST
4	58.5	35.7	453 1	SYS_ARCFU
5	58	35.4	301 1	YFE4_YEAST
6	57	35.2	180 1	SMCX_MOUSE
7	57	34.8	898 1	CIT21_HUMAN
8	57	35.2	1539 1	SMCY_HUMAN
9	57	35.2	1548 1	SMCY_MOUSE
10	57	35.2	1554 1	SMCX_MOUSE
11	57	35.2	1560 1	SMCX_MOUSE
12	54	32.9	222 1	T2M1_METJA
13	54	32.9	500 1	ERR2_HUMAN
14	53.5	32.6	232 1	ERR2_HUMAN
15	53	32.3	433 1	ERR2_RAT
16	53	32.3	458 1	ERR3_HUMAN
17	52	31.7	3140 1	POLG_PPVRA
18	52	31.7	3140 1	POLG_PPVSK

19	51.5	31.4	1043 1	EF3A_YEAST	P16521 saccharomyc
20	51	31.1	350 1	RRBB_RHSN	P55462 rhizobium s
21	51	31.1	390 1	YMF7_CAEBL	P34467 caenorhabdi
22	51	31.1	407 1	Y014_CAEBL	P34670 caenorhabdi
23	51	31.5	507 1	SEPI_MOUSE	P08197 mus musculu
24	51	31.1	845 1	KRUH_DROME	P08195 drosophila
25	50	30.5	116 1	ZTR6_HUMAN	Q00488 homo sapien
26	50	30.5	116 1	ZTR6_MOUSE	Q94b42 mus musculu
27	50	30.5	260 1	GRA_MOUSE	P11032 mus musculu
28	50	30.5	514 1	YUJ1_HAETN	P44744 hemophilus
29	50	30.9	548 1	SYEB_METJA	Q58508 methanococc
30	50	30.5	717 1	PRDD_HUMAN	Q94q73 homo sapien
31	50	30.5	733 1	YV34_HUMAN	Q94p66 homo sapien
32	50	30.9	993 1	ROXN_HUMAN	Q94g72 homo sapien
33	49.5	30.6	172 1	DCD_CLOTH	Q92n18 clostridium
34	49.5	30.2	961 1	BASO_MOUSE	Q35914 mus musculu
35	49	29.9	313 1	YV71_METJA	Q58471 methanococc
36	49	29.9	993 1	ROXN_HUMAN	Q94g72 homo sapien
37	48.5	29.6	809 1	LEF_BACAN	P15917 bacillus an
38	48	29.3	244 1	LPRA_MYCTU	Q11049 mycobacteri
39	48	29.3	295 1	MIRA_CAUCR	Q94635 caulobacter
40	48	29.3	332 1	PXCA_FUSNN	Q87970 fusobacteri
41	48	29.3	456 1	CN7A_MOUSE	P70453 mus musculu
42	48	29.6	457 1	IRP7_MOUSE	P70434 mus musculu
43	48	29.6	460 1	NU4M_SCYCA	Q79410 scyllorhinu
44	48	29.3	638 1	YM32_CAEBL	Q20680 caenorhabdi
45	48	29.3	646 1	NA95_HUMAN	Q901x6 homo sapien

ALIGNMENTS

RESULT 1
MAT3_RAT
ID MAT3_RAT STANDARD: PRT: 845 AA.
AC P43244; O35833;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Matrix 3.
GN MATR3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=91236771; PubMed=2033075;
RA Belgrader P., Dev R., Berzney R.;
RT "Molecular cloning of matrixin 3, A 125-kilodalton protein of the
RT nuclear matrix contains an extensive acidic domain.";
RT J. Biol. Chem. 266:9893-9899(1991).
RN [2]
RP REVISIONS.
RA Berzney R.;
RT Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
RT -1- FUNCTION: MAY PLAY A ROLE IN TRANSCRIPTION OR MAY INTERACT WITH
RT OTHER NUCLEAR MATRIX PROTEINS TO FORM THE INTERNAL FIBROGRAMULAR
RT NETWORK.
RT -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX.
RT -1- SIMILARITY: CONTRAINS 1 MATRIX-TYPE ZINC FINGER.
RT -1- SIMILARITY: CONTRAINS 2 RNA RECOGNITION MOTIFS (RNM).
RT This SWISS-PROT entry is copyright. It is produced through a collaboration
RT between the Swiss Institute of Bioinformatics and the EMBL outstation -
RT the European Bioinformatics Institute. There are no restrictions on its
RT use by non-profit institutions as long as its content is in no way
RT modified and this statement is not removed. Usage by and for commercial
RT entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
RT or send an email to license@isb-sib.ch).
DR EMBL: M63485; AAB63955.1; -;
DR InterPro: IPR000504; RNA_rec_mot.

```

DR InterPro: IPR003604; znf.U1.
DR InterPro: IPR000690; znf_mattrin.
DR Pfam: PF00076; trm. 2.
DR SMART: SM00360; RRM; 2.
DR SMART: SM00355; znf_C2H2; 1.
DR SMART: SM00451; znf_U1; 2.
DR PROSITE: PSS0102; RRM; 2.
DR PROSITE: PSS0030; RRM_RNP_1; FALSE_NEG.
DR PROSITE: PSS0171; 2F_MATIN; 1.
DR Nucleic acid binding protein; RNA-binding. Repeat; Zinc-finger.
FT DOMAIN 398 473
FT DOMAIN 496 571
FT DOMAIN 708 716
FT 2N_FING 799 830
FT NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT 2N_FING 799 830
FT NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 845 AA; 94446 MW; 5EABD05529171238 CRC64;

Alignment Scores:
Pred. No.: 0.154
Score: 63.00
Percent Similarity: 60.87%
Best Local Similarity: 39.13%
Query Match: 38.41%
DB: 1
Gaps: 0

US-09-513-151-3.COPY_1121-1210 (1-90) x MAT3_RAT (1-845)
QY 1 TGTGACCTGTGATGCATCATTCATTTGGGAGCGCATGAGGACACATATAATCC 60
Db 293 CysSerIleCysAspLeuProValHisSerAsnIlyGluTrpSerGlnHisIleAsnGly 312
QY 61 AAATCCAC 69
Db 313 AlaserrHis 315

RESULT 2
MAT3_HUMAN
ID MAT3_HUMAN STANDARD: PRT: 847 AA.
AC P43243; Q9UQ27; Q9UHW0;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Matrin 3.
GN MATR3 OR KIAA0723.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId:9606;
RX TISSUE=Brain;
RP SEQUENCE FROM N.A.
RC MEDLINE=9087487; PubMed=9872452;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RA "Prediction of the coding sequences of unidentified human genes. XI.
RA The complete sequences of 100 new cDNA clones from brain which code
RA for large proteins in vitro.";
RL DNA Res. 5:277-286(1998).

(2)
SEQUENCE FROM N.A.
RP TISSUE=Hypothalamus;
RX MEDLINE=20402571; PubMed=10931946;
RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Kong R., Ye M., Zhou J.,
RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
RA "Gene expression profiling in the human hypothalamus-pituitary-adrenal
RA axis and full-length cDNA cloning.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).

(3)
SEQUENCE OF 429-847 FROM N.A.
RX MEDLINE=9136771; PubMed=2033075;
RA Belgrader P., Dey R., Berezney R.;

```

RT	"Molecular cloning of matrin 3, A 125-kDoladon protein of the
RL	nuclear matrix contains an extensive acidic domain.";
CC	J. Biol. Chem. 266:9893-9899(1991).
CC	- FUNCTION: MAY PLAY A ROLE IN TRANSCRIPTION OR MAY INTERACT WITH
CC	OTHER NUCLEAR MATRIX PROTEINS TO FORM THE INTERNAL FIBROGRANULAR
CC	NETWORK.
CC	- SUBCELLULAR LOCATION: NUCLEAR MATRIX.
CC	- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
CC	- SIMILARITY: CONTAINS 1 MATRIN-TYPE ZINC FINGER.
CC	-----
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC	or send an email to license@sdb.ch).
CC	-----
DR	EMBL; AB018266; BAA34443.1; -
DR	EMBL; AF117236; AAFL7217.1; -
DR	EMBL; M63483; NOT ANNOTATED_CDS.
DR	Genew; HGNC:6912; MATR3.
DR	MIM; 604706; -
DR	InterPro: IPR000504; RNA_rec_mot.
DR	InterPro: IPR003604; znf_U1.
DR	InterPro: IPR000690; znf_matrin.
DR	Pfam; PF00076; rrm_2.
DR	SMART; SMO0360; RRM_2.
DR	SMART; SMO0355; znf_C2H2_1.
DR	SMART; SMO0451; znf_U1_2.
DR	PROSITE; PS50102; RRM_2.
DR	PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR	PROSITE; PS50171; zf_MATRIN; 1.
KW	Nuclear protein; RNA-binding; Repeat; Zinc-finger.
FT	DNAIN 398 473 RNA-BINDING (RRM) 1.
FT	DNAIN 496 571 RNA-BINDING (RRM) 2.
FT	DNAIN 710 718 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT	ZN_FING 801 832 MATRIN-TYPE.
FT	CONFICT 257 257 P -> S (IN REF. 2).
FT	CONFICT 274 274 P -> S (IN REF. 3).
FT	CONFICT 279 374 EDFFGLAPKPYHCSTGDDLPVSHNKESQHINGASHSRRC
FT	CONFICT 279 374 QLLLETPENKPNPDGTGGTDPPMLQOSTIPPAIGLAPP
FT	CONFICT 279 374 PPSFHGGAVSPRG -> VDVHWLTETVEIDICALYIGOV
FT	CONFICT 279 374 ILIRSGVNISMEOVTYVDASFLASTONGSLTFMTODQMVI
FT	CONFICT 279 374 HSCCSLIQIQHEFFMDLHPFIIGDQDLOR (IN REF.
FT	CONFICT 279 374 2).
FT	CONFICT 382 396 NLGGRPHNQKRIVET -> TCKDDLDCKRAEMKL (IN
FT	CONFICT 382 396 REF. 2).
FT	CONFICT 404 410 DFOGRKN -> GFSTEERT (IN REF. 2).
FT	CONFICT 429 431 LIL -> IPV (IN REF. 3).
FT	CONFICT 572 572 Y -> C (IN REF. 2 AND 3).
FT	CONFICT 691 691 G -> P (IN REF. 3).
FT	CONFICT 703 703 D -> H (IN REF. 3).
SO	SEQUENCE 847 AA; 94623 MW; 530AAA9214BC1611 CNC64;
Alignment Scores:	
Pred. No.:	0.154 Length: 847
Score:	63.00 Matches: 9
Percent Similarity:	60.87% Conservative: 5
Best Local Similarity:	39.13% Mismatches: 9
Query Match:	38.41% Indels: 0
DB:	Gaps: 1
US-09-1513-151-3_COPY_1121_1210 (1-90) x MAT3_HUMAN (1-847)	
OY	1 TTGTGCACCTCTGATCGAATCATCTTTGGAGTCGCAATGGAGGCCACATAAATC 60
Db	233 Cysseritiecysaspleuprovalnttssersnlysglittpserclnhistlleasnoly 312
Oy	61 AAATCCAC 69
Db	313 Alaserhis 315

RESULT 3
MOD5_YEAST STANDARD: PRT: 428 AA.
AC P07864: 012203;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE tRNA isopentenyltransferase (EC 2.5.1.8) (isopentenyl-diphosphate: DE tRNA isopentenyltransferase) (IPTase) (IPTase) (IPTase).
GN MOD5 OR Y0R274W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
ON NCBI_TaxID=4932;
RX MEDLINE-87172703; PubMed-3031457;
RA Najarian D., Dhanich M.E., Martin N.C., Hopper A.K.;
RT "DNA sequence and transcript mapping of MOD5: features of the 5' RT region which suggest two translational starts.";
RL Mol. Cell. Biol. 7:185-191(1987).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN-S288C;
RX MEDLINE-97051594; PubMed-8896271;
RA Cheret G., Bernardi A., Sor F.J.;
RT "DNA sequence analysis of the VP11-SNF2 region on chromosome XV of RT Saccharomyces cerevisiae.";
RL Yeast 12:1059-1064(1996).
[3]
RN ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.
RX MEDLINE-92052176; PubMed-1946403;
RA Slusher L.B., Gillman E.C., Martin N.C., Hopper A.K.;
RT "mRNA leader length and initiation codon context determine RT alternative AUG selection for the yeast gene MOD5.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9789-9793(1991).
[4]
RN ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.
RX MEDLINE-91203856; PubMed-1850093;
RA Gillman E.C., Slusher L.B., Martin N.C., Hopper A.K.;
RT "MOD5 translation initiation sites determine No-Isopentenyladenosine RT modification of mitochondrial and cytoplasmic tRNA.";
RL Mol. Cell. Biol. 11:2382-2390(1991).
[5]
RN ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.
RX MEDLINE-94187700; PubMed-8139535;
RA Boguta M., Hunter L.A., Shen W.C., Gillman E.C., Martin N.C.,
RT "Subcellular locations of MOD5 proteins: mapping of sequences RT sufficient for targeting to mitochondria and demonstration that RT mitochondrial and nuclear isoforms comingle in the cytosol.";
RL Mol. Cell. Biol. 14:2298-2306(1994).
[6]
RN FUNCTION: RESPONSIBLE FOR THE MODIFICATION OF A37 TO ISOPENTENYL A37 OF BOTH CYTOSOLIC AND MITOCHONDRIAL TRNAS.
CC -1 CATALYTIC ACTIVITY: Isopentenyl diphosphate + tRNA = diphosphate + CC tRNA containing 6-Isopentenyladenosine.
CC -1 PATHWAY: BIOSYNTHESIS OF THE MODIFIED BASE ISOPENTENYLADENOSINE CC IN TRNAS.
CC -1 SUBCELLULAR LOCATION: MITOCHONDRIAL, CYTOPLASMIC AND NUCLEAR.
CC -1 ALTERNATIVE PRODUCTS: TWO ISOFORMS MAY BE PRODUCED FROM THE USE OF CC ALTERNATIVE INITIATION CODONS IN THE SAME READING FRAME.
CC -1 SIMILARITY: BELONGS TO THE IPT TRANSFERASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration CC between the Swiss Institute of Bioinformatics and the EMBL Outstation- CC the European Bioinformatics Institute. There are no restrictions on its CC use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See http://www.isb-slb.ch/announce/ CC or send an email to license@isb-slb.ch).
CC EMBL; M15991; AAA34785.1; -

DR EMBL; X89633; CAA61780.1; -
DR EMBL; Z75182; CAA9499.1; -
DR PIR; A26717; A26717.
DR SGD; S0005800; MOD5.
DR InterPro: IPR002627; IPT.
DR Pfam; PF01715; IPT; 1.
DR ProDom; PD004674; IPT; 1.
DR TIGRfam; TIGR00174; miaA; 1.
KW Transferase; tRNA processing; ATP-binding; Alternative Initiation;
KW Mitochondrion; Nuclear protein.
FT CHAIN 1 428
FT
FT CHAIN 12 428
FT
FT INIT_MET 12 12
FT NP_BIND 21 28
FT DOMAIN 210 232
FT
FT CONFLICT 313 313
FT CONFLICT 375 375
SQ SEQUENCE 428 AA; 50236 MW; A956B17ABC05161F CRC64;

Alignment Scores:
Pred. No.: 0.616 Length: 428
Score: 59.00 Matches: 10
Percent Similarity: 57.14% Conservative: 10
Best Local Similarity: 28.57% Mismatches: 9
Query Match: 35.98% Indels: 6
DB: 1 Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x MOD5_YEAST (1-428)
QY 1 TGTGACCTCTGTG-----GATCGAATCATCATGCGGATCGCAATGG 42
|||:|||||:|
Db 375 CysAsnValCysArgAsnAlaAspCylLysAsnValAlaIleGlcLulysYrrTP 394
QY 43 GCAGGCGACATTAATCCCAATCCCATCTGACCACTGAAGAA 87
|||||:|||||:|
Db 395 LysIleHsLstLeuGlySerArgArgHisLysSerAenLulysArg 409

RESULT 4
SYS_ARCFU STANDARD: PRT: 453 AA.
AC 028244;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ser11-CRNA synthetase (EC 6.1.1.11) (Sertine-CRNA ligase) (SERTS).
GN SERS OR AF2035.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
ON NCBI_TaxID=2234;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE-98049343; PubMed-9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwyn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Klotzner E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadov P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -1 CATALYTIC ACTIVITY: ATP + L-serine + tRNA(Ser) = AMP + diphosphate
CC + L-seryl-tRNA(Ser).

```

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000962; AAB89219.1; -
DR HSSP: P34945; 1SES.
DR TIGR: AF2035; -
DR InterPro: IPR002106; AATRNA_Ligase1.
DR InterPro: IPR002314; tRNA-synt-2b.
DR InterPro: IPR002317; tRNA-synt-ser.
DR Pfam: PF00587; tRNA-synt-2b; 1.
DR Pfam: PF02403; SerYLtRNA.N; 1.
DR PRINTS: PR00981; TRNASYNTHSER.
DR TIGRfams: TIGR00414; serS; 1.
DR PROSITE: PS50862; AA_tRNA_Ligase_II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 453 AA; 52455 MW; 4BCE9EBA4F606D90 CRC64;

Alignment Scores:
Pred. No.: 0.739 Length: 453
Score: 58.50 Matches: 13
Percent Similarity: 62.07% Conservative: 5
Best Local Similarity: 44.83% Mismatches: 8
Query Match: 35.67% Indels: 3
DB: 1 Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x SYS_ARCFU (1-453)
QY 4 GACCTCTGTGATCATCATTCATTCGGGATCGGATCGGACGACATTAATCAAA 63
DB 27 AsplleValAspArgAlaIleGluLeuAspArgLysTrpArgGluLeuLys----- 44
QY 64 TCCCACTTGACCAACGACGAGAAAGA 90
DB 45 ---ArgValAsnGlnLeuArgLysArg 52

RESULT 5
YFE4_YEAST
ID YFE4_YEAST STANDARD; PRT; 301 AA.
AC P43558;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 33.5 kDa protein in SEC53-FET5 Intergenic region.
GN YFL044C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murkaml Y., Natou M., Haglawa H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Seda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae."
RL Nat. Genet. 10:261-268(1995).
CC -1- SIMILARITY: TO S.POMBE SPAC24C9.14.
CC -1- SIMILARITY: CONTAINS 1 OTU DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D50617; BAA09197.1; -
DR SGD: S0001850; YFL044C.
DR InterPro: IPR003323; OTU.
DR Pfam: PF02338; OTU; 1.
DR SMART: SM00355; ZnF_C2H2; 1.
DR PROSITE: PS50802; OTU; 1.
KW Hypothetical protein.
FT DOMAIN 109 OTU.
SQ SEQUENCE 301 AA; 33510 MW; 32ABA3CFEA4B5EB CRC64;

Alignment Scores:
Pred. No.: 0.86 Length: 301
Score: 58.00 Matches: 9
Percent Similarity: 60.00% Conservative: 6
Best Local Similarity: 36.00% Mismatches: 10
Query Match: 35.37% Indels: 0
DB: 1 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x YFE4_YEAST (1-301)
QY 1 TGTGACCTCTGTGATCATCATTCATTCGGGATCGGATCGGACGACATTAATCC 60
DB 272 CysAsnThrCysGlnMetThrPheValGlyGluArgGluValAlaArgHisAlaGluSer 291
QY 61 AATGCCACTTGAC 75
DB 292 ThrGlyHisValAsp 296

RESULT 6
SMCX_CRIGR
ID SMCX_CRIGR STANDARD; PRT; 180 AA.
AC P41228;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE SMCX protein (Xei169 protein) (Fragment).
GN SMCX OR Xei169.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95038834; PubMed=7951318;
RA Wu J., Salido E., Yen P., Mohandas T., Shapiro L.J.;
RT "The murine Xei169 gene escapes X-inactivation like its human
RT homologue."
RL Nat. Genet. 7:491-496(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L29564; AAA62383.1; -
FT NON_TER 1
FT NON_TER 180
SQ SEQUENCE 180 AA; 19964 MW; 89D2A3C50B4A703A CRC64;

Alignment Scores:
Pred. No.: 1.19 Length: 180
Score: 57.00 Matches: 9
Percent Similarity: 64.00% Conservative: 7
Best Local Similarity: 36.00% Mismatches: 3

```


RT "Prediction of the coding sequences of unidentified human genes. VI.
RT The coding sequences of 80 new genes (K1A0201-K1A0280) deduced by
RT analysis of cDNA clones from cell line KG-1 and brain."
RL DNA Res. 3:321-329(1996).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20319030; PubMed=10861003;
RA Shen P., Wang F., Underhill P.A., Franco C., Yang W.-H., Roxas A.,
RA Sung R., Lin A.A., Hyman R.W., Vollrath D., Davis R.W.,
RA Cavalli-Sforza L., Oefner P.J.;
RT "Population genetic implications from sequence variation in four Y
RT chromosome genes";
RL Proc. Natl. Acad. Sci. U.S.A. 97:7354-7359(2000);
RN [5]
RP SEQUENCE OF 1-457 FROM N.A. (ISOFORM 1).
RX MEDLINE=97224581; PubMed=9060413;
RA Aguilnik A.I., Bishop C.E., Lerner J.L., Solovay V.V.;
RT "Analysis of mutation rates in the SMCY/SMCX genes shows that
RT mammalian evolution is male driven."
RL Mamm. Genome 8:134-138(1997).
RN [6]
RP SEQUENCE OF 1340-1478 FROM N.A.
RA Poloumienko A., Blecher S.;
RT "Exon-intron structure of SMCX and SMCY genes in bovine and swine";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: May play a role in spermatogenesis. (By similarity).
CC -1- ALTERNATIVE PRODUCTS: At least 2 isoforms; 1 (shown here) and 2;
CC are produced by alternative splicing.
CC -1- SIMILARITY: CONTAINS 1 ARID DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 JMC DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 JMCN DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 PHD-TYPE ZINC FINGERS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: U52191; AAC50806.1; -
DR EMBL: D87072; BAA13241.1; -
DR EMBL: AF273841; AAG00951.1; -
DR EMBL: U52365; AAC51135.1; -
DR EMBL: AF134849; AAK27839.1; -
DR Gene: HGNC:11115; SMCY.
DR MIM: 426000; -
DR InterPro: IPR001606; ARID.
DR InterPro: IPR003347; TF_Jm1C.
DR InterPro: IPR003349; TF_Jm1N.
DR InterPro: IPR004198; ZnF_C5HC2.
DR InterPro: IPR001965; ZnF_PHD.
DR Pfam: PF00628; PHD; 4.
DR Pfam: PF02373; Jm1C; 2.
DR Pfam: PF02375; Jm1N; 1.
DR Pfam: PF02928; zf-C5HC2; 2.
DR SMART: SM00501; BRIGHT; 1.
DR SMART: SM00545; Jm1N; 1.
DR SMART: SM00249; PHD; 2.
DR PROSITE: PS01359; ZF_PHD_1; 2.
DR PROSITE: PS50016; ZF_PHD_2; 1.
KM Zinc-finger. Repeat. Alternative splicing.
FT DOMAIN 13 59 Jm1N.
FT ZN_FING 76 184 ARID.
FT DOMAIN 316 362 PHD-TYPE 1.
FT ZN_FING 491 607 Jm1C.
FT ZN_FING 1174 1235 PHD-TYPE 2.
FT VARSPLIC 118 174 MISSING (IN ISOFORM 2).
FT CONFLICT 327 327 D -> N (IN REF. 1) AND 3).
FT CONFLICT 1186 1186 V -> L (IN REF. 1).
FT CONFLICT 1285 1285 S -> F (IN REF. 1).

FT CONFLICT 1352 1352 P -> L (IN REF. 6).
FT CONFLICT 1391 1391 D -> G (IN REF. 1).
SQ SEQUENCE 1539 AA; 174072 MW; E58DAE374E3BD7AA CRC64;
Alignment Scores:
pred. No.: 1 37
Score: 57.00 Length: 1539
Percent Similarity: 64.008 Matches: 9
Best Local Similarity: 36.008 Conservative: 7
Query Match: 35.194 Mismatches: 3
DB: 1 Indels: 6
Gaps: 1
US-09-513-151-3_COPY_1121_1210 (1-90) x SMCY_HUMAN (1-1539)
QY 80 AGTTGTTCAAGTGGATTGTATGTCGCGCCATTCGGCATCCCAATGATG 21
Db 1221 Alatrptprlpunrprpshprlryspheleucys-----Proleucys 1234
QY 20 ATTGCATCAGAGAG 6
Db 1235 MetArgSerArgArg 1239
RESULT 9
SMCY_MOUSE
ID SMCY_MOUSE STANDARD; PRT; 1548 AA.
AC Q62240; Q9QYR9; Q9R040;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE SMCY protein (histocompatibility Y antigen) (H-Y).
GN SMCY OR HYA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RC STRAIN=129/SvJ;
RA Aguilnik A.I., Longepied G., Ty M.T., Bishop C.E., Mitchell M.J.;
RT "Mouse H-Y encoding SMCY gene and its X chromosomal homolog Smcx";
RL Mamm. Genome 10:926-929(1999).
RN [2]
RP SEQUENCE OF 23-586 FROM N.A.
RC STRAIN=BALB/C; TISSUE=Testis;
RC MEDLINE=95038738; PubMed=7524912;
RA Aguilnik A.I., Mitchell M.J., Lerner J.L., Woods D.R., Bishop C.E.;
RT "A mouse Y chromosome gene encoded by a region essential for
RT spermatogenesis and expression of male-specific minor
RT histocompatibility antigens";
RL Hum. Mol. Genet. 3:873-878(1994).
RN [3]
RP FUNCTION:
RX MEDLINE=95379952; PubMed=7544442;
RX Scott D.M., Ehrmann I.E., Ellis P.S., Bishop C.E., Aguilnik A.I.,
RA Simpson E., Mitchell M.J.;
RT "Identification of a mouse male-specific transplanted antigen, H-
RT Y";
RL Nature 376:695-698(1995).
RN [4]
RP FUNCTION:
RX MEDLINE=86146895; PubMed=3951555;
RA Burgoyne P.S., Levy E.R., McLaren A.;
RT "Spermatogenic failure in male mice lacking H-Y antigen";
RL Nature 320:170-172(1986).
CC -1- FUNCTION: May play a role in spermatogenesis.
CC -1- MISCELLANEOUS: SMCY encodes an H-Y epitope that is defined by the
CC octamer peptide TENGSKDI; since no similar peptide was found in
CC Smcx, it is presumably the genetic basis for the antigenic
CC difference between males and females that contributes toward a
CC tissue transplant rejection response.
CC -1- SIMILARITY: CONTAINS 1 ARID DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 JMC DOMAIN.


```

CC -1- SIMILARITY: CONTAINS 1 JMUN DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 PHD-TYPE ZINC FINGERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF127244; AAD53048.1;
DR EMBL: Z29652; CAA82760.1;
DR MGD: MGI:99780; Smcy.
DR InterPro: IPR001606; ARID.
DR InterPro: IPR003347; TE_JmJN.
DR InterPro: IPR003349; TE_JmJN.
DR InterPro: IPR004198; Znf_C5HC2.
DR InterPro: IPR001965; Znf_PHD.
DR Pfam: PF00628; PHD_4.
DR Pfam: PF01388; ARID_2.
DR Pfam: PF02373; JmJN_2.
DR Pfam: PF02375; JmJN_2.
DR Pfam: PF02928; zf-C5HC2; 1.
DR SMART: SM00501; BRIGHT; 1.
DR SMART: SM00545; JmJN; 1.
DR SMART: SM00249; PHD; 2.
DR PROSITE: PS01359; ZF-PHD_1; 2.
DR PROSITE: PS50016; ZF-PHD_2; 1.
KW Zinc-finger; Repeat.
FT DOMAIN 13 59 JMUN.
FT ZN_FING 76 184 ARID.
FT ZN_FING 325 371 PHD-TYPE 1.
FT DOMAIN 500 616 JmJN.
FT ZN_FING 1182 1243 PHD-TYPE 2.
FT CONFLICT 160 162 ILY -> SFT (IN REF. 2).
FT CONFLICT 167 167 F -> S (IN REF. 2).
FT CONFLICT 215 215 R -> K (IN REF. 2).
FT CONFLICT 583 586 VRIN -> STIK (IN REF. 2).
SQ SEQUENCE 1548 AA; 177016 MW; 44AAB94708EAA402 CRC64;

Alignment Scores:
Pred. No.: 1.37 Length: 1548
Score: 57.00 Matches: 9
Percent Similarity: 64.00% Conservative: 7
Best Local Similarity: 36.00% Mismatches: 3
Query Match: 35.19% Indels: 6
DB: Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x SMCY_MOUSE (1-1548)
QY 80 AGTGGTTCAGAGTGGATTGATTTATGCGCGCCCATTCGCATCCCAATGATG 21
Dy 1229 AlATrPrGtJutPrAsPThrlYsPhenLucys-----ProLucys 1242
QY 20 ATTGCATCAGAGAG 6
Dy 1243 MeLaYgSeArGaRg 1247

RESULT 10
SMCX_MOUSE STANDARD; PRT: 1554 AA
AC PA1230; Q8CVI4; Q9D0C3; Q9QVR8; Q9R039; O54995;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE SMCX protein (Xel69 protein).
GN SMCX OR XE169.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN-129/SyJ.
RX MEDLINE-99373259; PubMed-10441747;
RA Aguilnik A.I., Longepied G., Ty M.T., Bishop C.E., Mitchell M.J.;
RT "Mouse H-Y encoding Smcy gene and its X chromosomal homolog smcx."
RN Mamm. Genome 10:926-929(1999).
[2]
RP SEQUENCE OF 36-1068 FROM N.A.
RC STRAIN-BALB/C; TISSUE-Testis;
RX MEDLINE-95038739; PubMed-7951230;
RA Aguilnik A.I., Mitchell M.J., Matzel M.-G., Borsani G., Avner P.A.,
RT Lerner J.L., Bishop C.E.;
RT "A novel X gene with a widely transcribed Y-linked homologue escapes
RT X-inactivation in mouse and human."
RL Hum. Mol. Genet. 3:879-884(1994).
[3]
RP SEQUENCE OF 1206-1385 FROM N.A. (ISOFORM 1).
RX MEDLINE-95038834; PubMed-7951318;
RA Wu J., Salido E., Yen P., Mohandas T., Shapiro L.J.;
RT "The murine Xel69 gene escapes X-inactivation like its human
RT homologue."
RL Nat. Genet. 7:491-496(1994).
[4]
RP SEQUENCE OF 1-258 FROM N.A.
RC STRAIN-C57BL/6J; TISSUE-Embryo, and Small Intestine;
RX MEDLINE-21085660; PubMed-11217851;
RA Kawai J., Shunagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirini L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,
RA Brownstein M.J., Butt C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombert P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seiya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawai J., Kohsaki S.,
RA "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
[5]
RP SEQUENCE OF 1-39 FROM N.A.
RX MEDLINE-98389008; PubMed-9723615;
RA Jegalian K.G., Page D.C.;
RT "A proposed path by which genes common to mammalian X and Y
RT chromosomes evolve to become X inactivated."
RL Nature 394:776-780(1998).
[6]
RP STRAIN-C3H;
RC FUNCTION.
RX MEDLINE-95379952; PubMed-7544442;
RA Scott D.M., Ehrmann I.E., Ellis P.S., Bishop C.E., Aguilnik A.I.,
RT Simpson E., Mitchell M.J.;
RT "Identification of a mouse male-specific transplantation antigen,
RT H-Y."
RL Nature 376:695-698(1995).
-1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
-1- MISCELLANEOUS: Escapes X-inactivation.
CC -1- SIMILARITY: CONTAINS 1 ARID DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 JmJN DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 JMUN DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 PHD-TYPE ZINC FINGERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

GN smcX protein (aeros protein).
 GN SMCX OR XE169.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheraia; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 [1]

US-09-513-151-3_COPY_1121_1210 (1-90) x SMCX_HUMAN (1-1560)

OY		80	AGTTGGTCAAGAAGGCTATTGTGAATTTATGTGGCGCTGCCCATTCGGATCCGCACCAATGATG	21
			:::::::::::::::::::: :::	
Dd	1234	AlaIrrprIpLurpprAsphrrLySPrheLucyS-----ProlucyS		1247
OY		20	ATTGGATCACACAGAG	6
			::: :::	
Dd	1248	MetrGrSeIrArgrgrr	1252	

RESULT 12	
T2M1_METJA	
ID T2M1_METJA	STANDARD;
000001	PRT; 222 AA

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Type II restriction enzyme MjaI (EC 3.1.21.4) (Ecdonuclease MjaI)
 DE (R.MjaI)
 GN MJAI OR MJ0984.
 OS Methanococcus jannaschii.
 OC Archaea: Euryarchaeota: Methanococci: Methanococcales;
 OC Methanocaldococcaceae: Methanocaldococcus.
 CX NCBI_taxid=2190;

RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=9633799; PubMed=8688087;
RA Bult C.J., White O., Olsen G., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake O., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Getch C.T.,
RA Weierbach R., Kirkness E.F., Weinstock K.G., Merrick J.M., Gload A.,
RA Scott J.L., Geohegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utermack T.R., Kelley J.N., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts R.M., Hirst M.O., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.,
RT "complete genome sequence of the methanogenic archaeon, *Methanococcus*
RL *jannaschii*.", *Science* 273:1058-1073(1996).

RP CHARACTERIZATION.
PA Noren C.J., Roberts R.J., Pattl J., Byrd D.R., Morgan R.D.,
RT "Method for screening restriction endonucleases.",
RL Patent number WO91/1821, 11-MAR-1999.
CC
CC -1- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCE CTAG
CC AND CLEAVES AFTER C-1 (BY SIMILARITY).
CC
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give
CC specific double-stranded fragments with terminal 5'-phosphates
CC -1- SIMILARITY: STRON, TO M.THERMOFORMICUM MHZ1R.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL U67541; AAB98987.1; -.
DR REBASE 1221; MjA1.
DR TIGR; MJ0984; -.
KW Hydrolyase; Endonuclease; Nuclease; Restriction system
Complete proteome.
SO SEQUENCE 222 AA; 25962 MW; 9A2544208AB8AC02 CRC64

Alignment Scores:

Pred. No.:	3,52	Length:	222
Score:	54.00	Matches:	12

Percent Similarity:	72.738	Conservative:	4
---------------------	--------	---------------	---

Best Local Similarity:	54.55%	Mismatches:	6
Query Match:	32.93%	Indels:	0

DB:	1	Gaps:	0
DB:	1	Gaps:	0

US-09-513-151-3_COPY_1121_1210 (1-90) x T2M1_METJA (1-222)

[illegible]

QY	79	CTGAAG	84
Db	186	GLuLYs	187

```

RESULT 13
ERR2_HUMAN
ID ERR2_HUMAN
STANDARD;
PRT; 500 AA.

```

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Steroid hormone receptor ERR2 (Estrogen-related receptor, beta)
 DE ERRL-beta) (Estrogen receptor-like 2) (ERR beta-2).
 GN ESRB OR NR3B2 OR ERRB2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 XN NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
RC TISSUE-TESTIS;
RX MEDLINE=99173874; PubMed=10072763;
RA Chen F., Zhang Q., McDonald T., Davidoff M.J., Bailey W., Bai C.,
Liu Q., Caskey C.T.;
RT "Identification of two hERR2-related novel nuclear receptors utilizing
RT bioinformatics and inverse PCR."; *Gene* 228:101-109(1999).
RN 121
RP SEQUENCE OF 264-500 FROM N.A.
RA Rosen L., Medan A., Qin S., Baradaran L., Birditt B., Bloom S.,
RA Dots M., Dickhoff R., Fleetwood P., Harrison G., Kaur A., Medan A.,
RA Neebitt R., Tralcoff R., Hood L.;
RT "Sequencing of human chromosome 14q24.3 region.";

CC -1 SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1 SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1 SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC N3 SUBFAMILY.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/  
CC or send an email to license@slb.ch).  
CC -----  
DR EMBL; AF094517; AAC99409.1; -  
DR EMBL; AC008050; AAG17472.1; -  
DR HSSP; P03572; 1HCQ.  
DR TRANSFAC; T02766; -.  
DR
```

DR MI: 602167; -
DR InterPro: IPR000536; Hormone_rec_11g.
DR InterPro: IPR001723; Strdhmm_receptor.
DR InterPro: IPR001628; znf_C4steroid.
DR Pfam: PF00104; hormone_rec_1.
DR Pfam: PF00105; zf-C4; 1.

DR PRINT

```
DR Prodom; Pd0000035; Znf_C4steroid; 1.
DR SMART: SM00430; HOLT: 1
DR
```

```
DR SMART; SM00399; ZnF_C4; 1.
```

DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KE Recceptor; Transcription regulation; DNA-binding; Nuclear protein

receptor, transcription regulation, DNA binding, nuclear protein

FT	103	123	C4-TYPE.
ZN_FING			

```

SQ SEQUENCE 500 AA; 55619 MW; C219C84D914DF1C6 CRC64;

Alignment Scores:
Pred. No.: 3.72 Length: 500
Score: 54.00 Matches: 14
Percent Similarity: 57.14% Conservatave: 2
Best Local Similarity: 50.00% Mismatches: 6
Query Match: 32.93% Indels: 6
DB: 1 Gaps: 2

US-09-513-151-3_COPY_1121_1210 (1-90) x ERR2_HUMAN (1-500)

OY 1 TCGACCTGCTGATGATGCA-----ATCATCATGAGGATGCGAATGGCAGCGCAC 51
DB 244 CysAspLeuAlaAspArgGluLeuValValIleIleGly-----TrrPalatylshs 260
OY 52 ATAAATCCAAATCCCACTTGAAC 75
DB 261 IleProGlyPheSerSerLeuSer 268

RESULT 14
RADDC_BACHD STANDARD; PRT; 232 AA.
AC Q9K8H4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA repair protein radc homolog.
GN RADDC OR BH3032.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RA SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.,
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- FUNCTION: INVOLVED IN DNA REPAIR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RADDC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL: AP001517; BAB06751.1; -
DR InterPro: IPR000445; Hhm.
DR InterPro: IPR001405; RadC.
DR Pfam: PF00633; Hhm; 1.
DR ProDom: PD007415; RadC; 1.
DR TIGRFAMs: TIGR00608; radc; 1.
DR PROSITE: PS01302; RADDC; 1.
KW DNA repair; Complete proteome.
SQ SEQUENCE 232 AA; 26403 MW; 4542153B5C6010D2 CRC64;

Alignment Scores:
Pred. No.: 4.22 Length: 232
Score: 53.50 Matches: 11
Percent Similarity: 69.57% Conservatave: 5
Best Local Similarity: 47.83% Mismatches: 6
Query Match: 32.62% Indels: 1
DB: 1 Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x RADDC_BACHD (1-232)

OY 4 GACCTCTGATGATGATCATCATGAGGATGCGAATGGCAGCGCATATAATCCAA 63
DB 208 GluLeuLeuAspHisValIleIleGlyAspArgTrpPheValSer---LeuYsgLulys 226
OY 64 TCCCACTTG 72
DB 227 GlyHisLeu 229

RESULT 15
ERR2_RAT STANDARD; PRT; 433 AA.
AC P11475;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Steroid hormone receptor ERR2 (Estrogen-related receptor, beta) (ERR-
DE beta) (Estrogen receptor-like 2).
GN ESRB OR NR3B2 OR ESR2 OR ERR2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RA SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=88122546; PubMed=3267207;
RA Giguere V., Yang N., Segui P., Evans R.M.;
RT "Identification of a new class of steroid hormone receptors."
RL Nature 331:91-94(1988).
RN [2]
RP SHOWS THAT REF.1 ORIGINATES FROM RAT.
RX MEDLINE=99173874; PubMed=10072763;
RA Chen F., Zhang Q., McDonald T., Davidoff M.J., Bailey W., Bai C.,
RA Liu Q., Caskey C.T.;
RT "Identification of two hERR2-related novel nuclear receptors utilizing
RT bioinformatics and inverse PCR."
RL Gene 228:101-109(1999).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC -1- NR3 SUBFAMILY.
CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO ORIGINATE FROM HUMAN
CC BUT WAS LATER SHOWN (REF.2) TO BE DERIVED FROM RAT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL: X51417; CAA3779.1; -
DR PIR: B29345; B29345.
DR HSPD: P03372; IHC0.
DR InterPro: IPR000536; Hormone_rec_119.
DR InterPro: IPR001723; Sterhnm_receptor.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF00104; hormone_rec; 1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR00398; STRDHORMNER.
DR PRINTS: PR00047; STROIDFINGER.
DR ProDom: PD000035; Znf_C4steroid; 1.
DR SMART: SM00430; HOL1; 1.
DR SMART: SM00399; Znf_C4; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger.
DR DNA_BIND 103 168 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 103 123 C4-TYPE.
FT FT 139 163 C4-TYPE.
SQ SEQUENCE 433 AA; 48287 MW; EE2C4C5B2P9A3E13 CRC64;

```

Alignment Scores:

Pred. No.:	5.26	Length:	433
Score:	53.00	Matches:	11
Percent Similarity:	58.33%	Conservative:	3
Best Local Similarity:	45.83%	Mismatches:	10
Query Match:	32.32%	Indels:	0
DB:	1	Gaps:	0

US-09-513-151-3_COPY_1121_1210 (1-90) x ERR2_RAT (1-433)

QY	1	TGTGACCTCTGTGATCGATCATCATTTGGGGATCGCGAATGGCGCGCACATAAAATCC	60
Db	244	CysaspLeuAlaaspArgGluLeuValPheLeuIleSerTrpAlaLysHisIleProGly	263
QY	61	AAATCCCACTTG	72
Db	264	PheSerAsnLeu	267

Search completed: April 21, 2003, 18:49:20
Job time : 4.07696 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 21, 2003, 16:30:20 ; Search time 3.16753 Seconds

(without alignments)
7572.189 Million cell updates/sec

Title: us-09-513-151-3_copy_1121_1210

Perfect score: 164

Sequence: 1 TGTGACCTCTGTGATGCAAT.....TGAACCACTGAGAGAAAGA 90

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```

-MODEL=frame_n2p.model -DEV=xlh
-O=/cgn2_1/USPRO.spool/US09513151/runat.15042003.141143.26358/app.query.fasta_1.2446
-DB=A.GeneSeq.101002 -QFMT=fastan -SUFFIX=n2p_rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pcpt -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09513151 -ECGN.1_1_84=/runat.15042003.141143.26358 -NCPU=3
-NO.XLPHY -NO.MMAP -LARGOQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database : A.GeneSeq.101002:*

```

1: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	164	100.0	411	23	Novel human protei
2	64	39.0	302	21	Human ORFX ORF1724
3	63	38.4	617	23	Human ovarian anti
4	63	38.4	864	22	Novel human diagno
5	61	37.2	314	22	Drosophila melanog
6	60	36.6	60	23	Human ORFX protein
7	60	36.6	430	20	Amino acid sequenc
8	58	35.8	72	21	Arabidopsis thalia
9	58	35.8	72	21	Arabidopsis thalia
10	58	35.8	121	21	Arabidopsis thalia
11	58	35.4	578	22	Arabidopsis thalia
12	58	35.4	884	22	Drosophila melanog
13	58	35.4	935	22	Drosophila melanog
14	57.5	35.1	935	22	Breast and ovarian
15	57	34.8	737	21	Human transcriptio
16	57	34.8	765	22	Human protein sequ
17	57	34.8	786	22	Human polypeptide
18	57	34.8	842	22	Human polypeptide
19	57	34.8	898	22	Human polypeptide
20	57	35.2	1539	18	Human SMC2 protein
21	56	34.1	269	22	Human protein sequ
22	56	34.1	346	22	Drosophila melanog
23	55.5	33.8	539	22	Drosophila melanog
24	55	33.5	652	22	Drosophila melanog
25	55	33.5	789	22	Drosophila melanog
26	54	32.9	123	21	Human secreted pro
27	54	32.9	187	21	Human secreted pro
28	54	32.9	433	21	Human secreted pro
29	54	32.9	500	20	Human nuclear rece
30	54	32.9	500	20	Human nuclear rece
31	54	32.9	508	20	Human nuclear rece
32	54	32.9	508	20	Human nuclear rece
33	53.5	32.6	238	22	Human novel secret
34	53.5	32.6	1424	22	Human polypeptide
35	53.5	32.6	1464	22	Human polypeptide
36	53	32.3	340	22	Human polypeptide
37	53	32.3	341	22	Human polypeptide
38	53	32.3	418	20	Human nuclear rece
39	53	32.3	418	20	Human nuclear rece
40	53	32.3	433	9	Sequence of human
41	53	32.3	433	21	Human ERbeta prot
42	53	32.3	435	19	Steroid hormone re
43	53	32.3	435	21	Human oestrogen re
44	53	32.3	458	20	Human nuclear rece
45	53	32.3	458	21	Human ERgamma pro

ALIGNMENTS

RESULT 1
ABB97295 standard; Protein: 411 AA.

27-JUN-2002 (first entry)

Novel human protein SEQ ID NO: 563.

Human; antihaemetic; vulnary; antinflammatory; immunomodulator;
antifertility; cerebroprotective; cytosolic; rheumatic; gene therapy;
neuroprotective; antiparkinsonian; protein therapy; EST;
expressed sequence tag.

XX Homo sapiens.
XX OS
XX PN W0200222660-A2.
XX

XX	21-MAR-2002.
PF	10-SEP-2001; 2001MO-US26015.
XX	
PR	11-SEP-2000; 2000US-0659671.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QH, Ren F;
PI	Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX	
DR	WPI: 2002-292408/33.
DR	N-PDSB; ABN32481.
XX	
PT	An isolated polynucleotide for treating diseases associated with its
PT	encoded polypeptide such as cancer and multiple sclerosis -
XX	
PS	Example 2; SEQ ID NO 563; 509bp; English.
XX	
CC	The present invention provides the protein and coding sequences of 444
CC	novel human proteins. These were isolated from expressed sequences tags
CC	(ESTs). They can be used to stimulate cell growth, to regulate
CC	hematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC	e.g. in burn treatment, to regulate the immune system e.g. to treat
CC	multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC	fertility, to regulate haemostasis or thrombolysis e.g. to treat
CC	stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC	e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC	Parkinson's disease. The present sequence is a protein of the invention.
XX	
SQ	Sequence 411 AA:
	Alignment Scores:
	Pred. No.: 3,56e-18 Length: 411
	Score: 164.00 Matches: 30
	Percent Similarity: 100.00% Conservative: 0
	Best Local Similarity: 100.00% Mismatches: 0
	Query Match: 100.00% Indels: 0
	DB: 23 Gaps: 0
US-09-513-151-3_COPY_1121_1210 (1-90) x ABB97295 (1-411)	
OY	1 TGTCACCTGTGATCGAATCATATTGGGATGCCAAATGGGCAGGCACATAAATCC 60
Db	341 CysaspLeucCysAspArgIleIleIleIleGlyAspArgGlnTrrAlaIahHisIleLysSer 360
OY	61 AAATCCCACTGAACAACAAGTAAGAAGA 90
Db	361 LysSerHisLeuAengInLeuLysLysArg 370
RESULT 2	
ID	AAB41960 standard; Protein: 302 AA.
AA	AAB41960
XX	
CC	AAB41960;
XX	
DT	08-FEB-2001 (first entry)
DE	Human ORFX ORE1724 polypeptide sequence SEQ ID NO:3448.
XX	
KM	Human: open reading frame: ORFX; detection: cytostatic; hepatotropic;
KM	vulnerable; antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
KM	anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiantic;
KM	immunostimulant; thrombolytic; coagulant; vasodilator; antidiabetic;
KM	hypotensive; demagogical; immunosuppressive; antiinflammatory;
KM	antiviral; antibacterial; antifungal; antirheumatic; antihypoid;
KM	antiangiemic; gene therapy; cancer; proliferative disorder; hypertension;
KM	neurodegenerative disorder; osteoarthritis; graft vs host disease;
KM	cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KM	cholesterol ester storage; systemic lupus erythematosus; infection;
KM	severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KM	allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KM	bone damage; cartilage damage; antiinflammatory disease; coagulation;

```

KM thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
FN MO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-USO8621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
PI Shinkets RA, Leach M;
DR WPI: 2000-602362/57.
XX N-PSDB; AAC76169.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancer, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
PT
PS Claim 11; Page 2624; 5507pp; English.
XX
AC AACT74446 to AACT77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antiparasitic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticoagulant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiatic; thrombolytic; coagulant; vasotrophic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiflammatory; antibacterial; antiviral; antifungal; antineumatic;
CC antihypoid; and antihaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hyperension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 302 AA;
XX
Alignment Scores:
Pred. No.: 0 12 Length: 302
Score: 64.00 Matches: 10
Percent Similarity: 65.52% Conservative: 9
Best Local Similarity: 34.48% Mismatches: 10
Query Match: 39.02% Indels: 0
DB: 21 Gaps: 0
US-09-513-151-3_COPY_1121_1210 (1-90) x AABA1960 (1-302)
QY 1 TGTGACCTCTGTGATGCAGTCAATCATTCATTGGGAGATGCCGAATGGCAGCGCATMAATACC 60
Db 83 CysLysValcysScySalAlaLeuLeIILSeSLnserLnLynSLeuAlaHtSYrGlnSer 102
QY 61 AAATCCCATTTGAACAACAACATGAGAA 87
Db 103 LysLysHtsAlaIsalnsLysVallysgArg 111
RESULT 3
ABP41976
ID ABP41976 standard; Protein: 617 AA.
```

XX AC ABP41976;
 XX XX 22-AUG-2002 (first entry)
 DT XX
 XX XX Human ovarian antigen HAO63, SEQ ID NO:3108.
 DE XX
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antinflammatory; gynaecological; reproductive.
 KW XX
 OS Homo sapiens.
 XX OS
 XX WO200200677-A1.
 PN 03-JAN-2002.
 PD 07-JUN-2001; 2001WO-US18569.
 XX 07-JUN-2001; 2000US-209467P.
 PF 07-JUN-2000; 2000US-209467P.
 XX
 PR (HOMA-) HUMAN GENOME SCI INC.
 XX
 PA Birse CE, Rosen CA;
 XX
 PI WPI: 2002-147878/19.
 DR N-PSDB: ABQ55053.
 XX
 DR Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 PS Claim 11; SEQ ID NO 3108; 29222pp; English.
 XX
 XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54331-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis), and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 617 AA;

Alignment Scores:
 Pred. No.: 0.203
 Score: 63.00
 Percent Similarity: 60.87%
 Best Local Similarity: 39.13%
 Query Match: 38.41%
 DB: 23
 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x ABP41976 (1-617)

QY 1 TGTGACCTCTGTGATCGAATCATTCATGGGATCCGGAATGGCAGCCGCACTAAATCC 60
 ||| ::|||
 Db 69 CysSerIleCysAspLeuProValHisSerAsnLysGluTrpSerClnHisIleAsnGly 88
 ||| |||||

QY 61 AATCCAC 69
 |||||

Db 89 AlaserHis 91

RESULT 4
 ABG18701
 ID ABG18701 standard: Protein; 864 AA.
 XX
 AC ABG18701;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #18692.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 OS WO200175067-A2.
 PN 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 PF 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 PR (HYSE-) HYSEQ INC.
 XX
 PA Drmanac RT, Liu C, Tang YT;
 XX
 PI WPI: 2001-639362/73.
 DR N-PSDB: AAS82888.
 XX
 DR New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID NO 49060; 103pp; English.
 XX
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and

CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORP proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 AA;

Alignment Scores:
Pred. No.: 0 402 Length: 60
Score: 60.00 Matches: 11
Percent Similarity: 58.33% Conservative: 3
Best Local Similarity: 45.83% Mismatches: 10
Query Match: 36.59% Indels: 0
DB: 23 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x ABP04585 (1-60)

QY 18 AATCATCATTTGGGATCGCATGCGACACATAAATCCAAATCCACTTGAACA 77
||| ||||| ::||| ||||| :: ||||| |||||
Db 24 AAsArGlsrPrProThrArGAlaGlySerProThrArGArProLeuPrThGluPro 43

QY 78 ACAGCAAGAAAG 89
||| |||||
Db 44 GlyProGluLys 47

RESULT 7

ID AAY02526 standard; Protein: 430 AA.

XX AAY02526;

DT 15-JUL-1999 (first entry)

XX Amino acid sequence of the GRO-1 protein.

XX gro-1 operon; gro-1 gene; gop-1 gene; gop-2 gene; gop-2 gene;

KW hap-1 gene; cancer; aging; longevity; tumour formation;

KW physiological clock.

XX Caenorhabditis elegans.

OS Caenorhabditis elegans.

XX WO910482-A1.

PN 04-MAR-1999.

PD 20-AUG-1998; 98MO-CA00803.

PF 25-AUG-1997; 97CA-2210251.

XX (UYMC-) UNIV MCGILL.

PA Barnes T, Hekimi S, Lakowski B, Lemieux J;

XX WPI; 1999-190615/16.

DR N-PSDB; AAX36072.

XX Molecular identity of the gro-1 gene - useful for cancer diagnosis

PT and/or prognosis, and where compounds affecting encoded proteins are

PT useful for enhancing longevity of a host and inhibiting tumour

PT formation

XX C1a1m 11; Page 47-48; 93pp; English.
PS The present sequence represents the GRO-1 protein. The specification
XX describes the five genes of the Caenorhabditis elegans gro-1 operon
CC (AAX36071). The operon contains the gro-1 gene (AAX36072), the
CC gro-1 gene (AAX36074), the gop-2 gene (AAX36075), the gop-2 gene
CC (AAX36075), and the hap-1 gene (AAX36077). The gro-1 gene can be used in
CC a method for the diagnosis and/or prognosis of cancer in a patient.
CC Transgenic mice containing a gene knock-out of a murine gene homologous
CC of the gro-1 gene are useful as models of aging and cancer. The
CC proteins encoded by the genes are useful for identifying compounds
CC that affect the enzymatic activity of these proteins, in order to
CC enhance longevity of a host and inhibit tumour formation. The gro-1
CC gene, together with the gop-1, gop-2, gop-3 and hap-1 genes enables
CC study of a physiological clock.
XX
SQ Sequence 430 AA;

Alignment Scores:
Pred. No.: 0 593 Length: 430
Score: 60.00 Matches: 9
Percent Similarity: 53.33% Conservative: 7
Best Local Similarity: 30.00% Mismatches: 14
Query Match: 36.59% Indels: 0
DB: 20 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x AAY02526 (1-430)

QY 1 TGTGACCTGTGTGATCGATCATTTGGGATCGCATGCGACACATAAATCC 60
||| ||||| ::||| ||||| :: ||||| |||||
Db 394 CysGluIleCysAsnIleSerMetThrGlyLysAspAsnTrpGlnLysIleAspGly 413

QY 61 AATGCCACTTGACACCAACTGACAAAGA 90
||| |||||
Db 414 LysLysHisLysHisHisAlaLysGlnLys 423

RESULT 8

ID AAG56792 standard; Protein: 72 AA.

XX AAG56792;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 73087.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PN 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.

PF 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135153.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142905.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143524.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155436.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158322.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.

PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 0.893 Length: 72
Score: 58.00 Matches: 12
Percent Similarity: 51.61% Conservative: 4
Best Local Similarity: 38.71% Mismatches: 9
Query Match: 35.80% Indels: 6
DB: 21 Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x AAG56792 (1-72)

OY 80 AGTGGCTCAAGTGGGATTTGATGTCGCGCTCCCATTCGCCATCCCAATGATG 21
DB 27 SetrTpaRgLeuRtpgIuValPProPhetRPyCysValHshIsaspGlnIleProLeuGly 46

OY 20 ATT-----CGATCAGAGG 6
DB 47 llelleleuLeuIleargGluArgSerGlnArg 57

RESULT 9
AAG60375
ID AAG60375 standard; Protein: 7? AA.
XX AAG60375;
AC
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 78196.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132653.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135141.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.

```

PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158393.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

```

```

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 0.893
Score: 58.00
Percent Similarity: 51.61%
Best Local Similarity: 38.71%
Query Match: 35.80%
DB: 21

US-09-513-151-3_COPY_1121_1210 (1-90) x AAG60375 (1-72)

OY 80 AGTGGTCAAGCGCATTTGATGTGCGCTGCCCATTCGGCATCCCATGANG 21
    ||||| |||:||||| ||| ||| |||:|||||
Db 27 SerTTPaRgLeuTPrGluValProPheTPrCySValHisHisApsInIleProLeuGly 46
OY 20 ATT-----CGATCAGAGG 6
    ||| ||||| ||||| ||||| ||||| |||||
Db 47 IleIleGluLeuLeuArgGluArgSerGlnArg 57

RESULT 10
AAG56791
ID AAG56791 standard; Protein; 121 AA.
XX
AC AAG56791;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 73086.
XX
KW protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.

```

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 13-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135153.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138447.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148665.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

```
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 0 99
Score: 58.00
Percent Similarity: 51.61%
Best Local Similarity: 38.71%
Query Match: 35.80%
DB: 21
Matches: 121
Conservative: 4
Mismatch: 6
Indels: 9
Gaps: 1
```

```
US-09-513-151-3_COPY_1121_1210 (1-90) x AAG56791 (1-121)
QY 80 AGTTGGTCAAGTGGATTGATTTATGTCGCTGCCATTCGCGATCCCAATGATG 21
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 76 SerTrrpArgLeuTrpGluValProPhetrrpCysValHisHisaspGlnIleProLeucGly 95
QY 20 ATT-----CGATCAGACAGC 6
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 96 IleIleGluLeuIleArgGluArgSerGlnArg 106
```

```
RESULT 11
AAG60374
ID AAG60374 standard; Protein; 121 AA.
XX
AC AAG60374;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 78195.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
```

```
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134421.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
```

```

PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149428.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157513.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.

```

```

PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159328.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 0.99
Score: 58.00
Percent Similarity: 51.61%
Best Local Similarity: 38.71%
Query Match: 35.80%
DB: 21

US-09-513-151-3_COPY_1121_1210 (1-90) x AAG60374 (1-121)

QY 80 AGTGTGATCAAGTGGATTGATTTATGTCGCTGCCCATTCGCAATCCCAATGATG 21
    ||||| |||:|:| || ||| |||:|:|
DB 76 SERTTPARGLEUTPGLVALPPOBHERPCYSVALHSHLSAASGLNLEPROLEUGY 95
    ||| |-----CGATCAGAGG 6
QY 20 ATT-----
DB 96 ILEILGLULEUARGSLNARGSERGINARY 106
    ||| ||||| |||||

RESULT 12
ABB58393
ID ABB58393 standard; Protein; 578 AA.
XX
AC ABB58393;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 1971.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
    pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX

```


DR WPI: 2001-656860/75.
DR N-PSDB; ABL02496.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 1971; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (AB101840-AB116175).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 578 AA:

Alignment Scores:
Pred. No.: 1.35 Length: 578
Score: 58.00 Matches: 11
Percent Similarity: 62.50% Conservative: 4
Best Local Similarity: 45.83% Mismatches: 7
Query Match: 35.37% Indels: 2
DB: 22 Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x ABB58393 (1-578)
OY 1 TGTGACCTCTGTGATGCATCATTCATGGGATCGGACGACCATTAATCC 60
DB 315 CysGluLeucCysGlyArgMetPheSerSerArgAspGluTrpSerIleHis-----Ala 332
OY 61 AATCCCACTTG 72
DB 333 LysSerHisIleu 336

RESULT 13
ABB61358
ID ABB61358 standard; Protein; 884 AA.
XX
AC ABB61358;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 10866.
XX
KM Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI: 2001-656860/75.
DR N-PSDB; ABL05461.
PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 10866; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (AB101840-AB116175).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 884 AA:

Alignment Scores:
Pred. No.: 1.47 Length: 884
Score: 58.00 Matches: 10
Percent Similarity: 53.57% Conservative: 5
Best Local Similarity: 35.71% Mismatches: 13
Query Match: 35.37% Indels: 0
DB: 22 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x ABB61358 (1-884)
OY 1 TGTGACCTCTGTGATGCATCATTCATGGGATCGGACGACCATTAATCC 60
DB 255 CysGluLeucCysAspAlaThrCysThrGlyThrAspAlaIleValAlaGly 274
OY 61 AATCCCACTTGAAACCACTGAG 84
DB 275 AlaIleHisGluIleValValIlys 282

RESULT 14
ABB58283
ID ABB58283 standard; Protein; 935 AA.
XX
AC ABB58283;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 1641.
XX
KM Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI: 2001-656860/75.
DR N-PSDB; ABL02386.
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 1641; 21pp + Sequence Listing; English.

[illegible]

```

PS      Claim 11: Page 1193-1196; 1299pp; English.
XX
XX      Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC      proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC      associated with breast and ovarian cancer. Included in the invention
CC      sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC      isolation and characterisation of the DNA and protein sequences of t
CC      invention. The breast and ovarian cancer associated DNA, protein, ag
CC      or antagonist sequences exhibit cytostatic; immunosuppressive;
CC      neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC      antidiabetic; antiinflammatory; antitumor; vulnerable; anticonvulsant
CC      antibacterial; antifungal; antiparasitic and cardiant activity. The
CC      polynucleotide and protein sequences are used in the diagnosis of ca
CC      particularly breast and ovarian cancer. The nucleic acid sequences,
CC      proteins, agonists and agonists may also be used in the diagnosis,
CC      prevention and treatment of immune disorders e.g. Addison's disease,
CC      allergies, autoimmune hemolytic anaemia, autoimmune thyroiditis,
CC      diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC      arthritis and ulcerative colitis; cardiovascular disorders such as
CC      myocardial ischaemias; wound healing; neurological diseases such as
CC      cerebral anoxia and epilepsy; and infectious diseases.
CC
XX      SQ      Sequence      737 AA:

Alignment Scores:
Pred. No.:      2.07      Length:      737
Score:      57.00      Matches:      11
Percent Similarity:      53.57%      Conservative:      4
Best local Similarity:      39.25%      Mismatches:      13
Query Match:      34.75%      Indels:      0
DB:      21      Gaps:      0

US-09-513-151-3_COPY_1121_1210 (1-90) x AAB59023 (1-737)
OY      1      TGTGACCTCTGTATCGATTCATCTATTTGGGATCGCGAATGGGAGCGCATTAATATCC      60
      |||::||| ||| ||| |||
DB      47      CyselIuencuSyaspalserCysThrGlylaaspalatyralaalaHisIleargly      66
OY      61      AATATCCACTTGAACCACTGAAG      84
      ||| ::|||
DB      67      AlalysHisGlnlyValVallyls      74

Search completed: April 21, 2003, 18:48:18
Job time : 5.16753 secs

```



```

Db      21 ArgThrLeuProLeuValIleLeuGlyAlaThrGlyThrGlyLysSerThrLeuAla 40
OY      131 TTGAGCTAGCCAGCGGCTCGGAGATGCTACGGCTGACTCCATGCAAGTTAT 190
        |||
Db      41 LeuGlnLeuGlnArgLeuGlnGlyGlyIleValSerIleAspSerMetGlnValTyr 60
OY      191 GAAGGCTTAGACATCATCACCAACAGTTTCGCCCAAGAGGAGAAATCTGCCGAC 250
        |||
Db      61 GlnGlyLeuAspIleIleThrAsnLysValSerIleGlnIleGlnArgIleCysArgHis 80
OY      251 CACATGATCAGCTTTGTGATCTCTGTGACCAATATACAGTGTGAGCTTCAGAAAT 310
        |||
Db      81 HisMetIleSerPheValAspProLeuValThrAsnThrThrValValAspPheArgAsn 100
OY      311 AGAGCAACTGCTGTGATTGAA-----GATATATTGCCCGA 346
        |||
Db      101 ArgAlaThrAlaLeuIleSerLeuGlyLysAlaAlaIleGlyPheAspIlePheAlaArg 120
OY      347 GACAAATTCCTATTGTGTGGGAGGAGAAACAATTATACATTTGCAATCTGCTCGAAA 406
        |||
Db      121 AspLysIleProIleValIleGlyGlyThrAsnTyrTrpIleGlnSerLeuLeuTrpLys 140
OY      407 GTTCTGTCAATACCAAGCCCAAGAGATGGGACTGAGAAAGTGAATGACCCGAAAGTG 466
        |||
Db      141 ValLeuValAsnThrLysProGlnGlnMetGlyThrGlnLysValIleAspArgLysVal 160
OY      467 GAGTTTAAAAAGAGAGTGTCTTTGTACTTCACAAAGCCTTAAGCCAGGTGGAGCCGAA 526
        |||
Db      161 GlnLeuGlnLysGlnAspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGln 180
OY      527 ATGGCTGCAGGCTGCATCCACATGACAAAGCAAGTGGCCAGGACTTGCAGAGTTT 586
        |||
Db      181 MetAlaIleLysLeuHisProHisAspLysArgLysValAlaArgSerLeuGlnValPhe 200
OY      587 GAAGAAACAGGAATCTCTCATAGTAATTTCTCCATGCTACATACGGAAGAAGTGTGT 646
        |||
Db      201 GlnGlnThrGlyIleSerHisSerGlnPheLeuHisArgGlnHisThrGlnGlnGly 220
OY      647 GGTCCCTTTGGAGGCTCTCTGGAAGTTCTTAACCTTGCACTTTGGCTTCATGCTGAC 706
        |||
Db      221 GlyProLeuGlyLysProLeuLysPheSerAsnProCysIleLeuTrpLeuHisAlaAsp 240
OY      707 CAGGCACTTAGATGAGCCCTTGATTAAGAGGCTGATACATGCTTGGCTGGGCTC 766
        |||
Db      241 GlnAlaValLeuAspGlnValArgLeuAspLysArgValAlaAspMetLeuAlaIleGlyLeu 260
OY      767 TTGGAGAGAACTAAGAGATTTTTCACAGACGCTATATACAGAAATGTTCGGAATAATAGC 826
        |||
Db      261 LeuGlnGlnLeuArgAspPheHisArgArgTyrAsnGlnLysAsnValSerGlnAsnSer 280
OY      827 CAGGACTATACATGATGATCTTCATCAATTTGGCTTCAAGGAATTTACAGAGTACTG 886
        |||
Db      281 GlnAspLysGlnHisGlyIlePheGlnSerIleGlyPheLysGlnPheHisGlnTyrLeu 300
OY      887 ATCACTAGAGGAAATGACACTGAGAGACTAGTACAGCTTCTTAAGAAGAAGA----- 940
        |||
Db      301 IleThrGlnGlnLysCysThrIleGlnIleThrSerAsnGlnLeuLeuLysGlyIleGln 320
OY      940 ----- 940
        |||
Db      321 AlaLeuLysGlnValThrLysArgTyrAlaArgLysGlnAsnArgTrpValLysAsnArg 340
OY      941 -----CCTGTGCCATTGTGCCCGCTGCTATAGGCTTAGAGGATCTGATGCTC 988
        |||
Db      341 PheLeuSerArgProGlyProIleValAlaProValIleGlyIleGlnGlnValSerAspVal 360
OY      989 TCGAAGTGGAGAGAGTCTGTCTTGAACCTGCTTGAATCGTGAATTTTCATCAG 1048
        |||
Db      361 SerLysTrpGlnGlnSerValLeuGlnProAlaLeuGlnIleValAlaGlnSerPheIleGln 380
OY      1049 GGGCACAAGCCTACAGCCACTCCCAATAAAGATGCCATACATGAAGCTGGAACACAGAGA 1108
        |||
Db      381 GlyHisLysProThrAlaThrProIleLysMetProTyrAsnGlnIleGlnValAsnLysArg 400

```

```

OY      1109 AGTTATCACCTGTGTACCTCTGTGATCGAATCATCATTTGGGATGCCAATGGCGACG 1168
        |||
Db      401 SerTyrHisLeuLysAspLysCysAspArgIleIleIleGlyAspArgGlnTrpAlaAla 420
OY      1169 CACATAAATTCAAATCCACTTGAACCACTGAAAGAAAGAGAGATGGACTAGAT 1228
        |||
Db      421 HisIleLysSerLysSerHisLeuAsnGlnLeuLysLysArgValGlnLeuAspSerAsp 440
OY      1229 GCTGTCAACACATAGAACTCAGAGTGTTCGCCAGACTATACAAAGAACCTAAAGG 1288
        |||
Db      441 AlaValAsnThrIleGlnSerGlnSerValSerProAspHisAsnLysGlnProLysGln 460
OY      1289 AAGGATCCCGAGGAGAGATGATCAAGACTGAATGACGCTT 1333
        |||
Db      461 LysGlySerProGlyGlnAsnAspGlnIleuLysCysSerVal 475

RESULT 2
US-60-453-050-7968
; Sequence 7968, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7968
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-7968

Alignment Scores:
Pred. No.: 1,02e-172 Length: 475
Score: 2234.00 Matches: 438
Percent Similarity: 92.42% Conservative: 1
Best Local Similarity: 92.21% Mismatches: 2
Query Match: 62.49% Indels: 34
DB: 7 Gaps: 2

US-09-513-151-3 (1-2041) x US-60-453-050-7968 (1-475)
OY      11 ATGGCGTCCGTGGCGGCTGCAGCAGCACTTCTGTGGCAGTGGGCTCAGAGGCGCTGCAA 70
        |||
Db      1 MetAlaSerValAlaAlaIleAlaArgAlaValProValIleSerGlyLeuArgGlyLeuGln 20
OY      71 CGGACCTACTCTTGTAGTGATTTCTGGGGCCACGGGCAATCCAGGCTGGCG 130
        |||
Db      21 ArgThrLeuProLeuValIleLeuGlyAlaThrGlyThrGlyLysSerThrLeuAla 40
OY      131 TTGAGCTAGCCAGCGGCTCGGAGATGCTACGGCTGACTCCATGCAAGTTAT 190
        |||
Db      41 LeuGlnLeuGlnArgLeuGlnGlyGlyIleValSerIleAspSerMetGlnValTyr 60
OY      191 GAAGGCTTAGACATCATCACCAACAGTTTCGCCCAAGAGGAGAAATCTGCCGAC 250
        |||
Db      61 GlnGlyLeuAspIleIleThrAsnLysValSerIleGlnIleGlnArgIleCysArgHis 80
OY      251 CACATGATCAGCTTTGTGATCTCTGTGACCAATATACAGTGTGAGCTTCAGAAAT 310
        |||
Db      81 HisMetIleSerPheValAspProLeuValThrAsnThrThrValValAspPheArgAsn 100
OY      311 AGAGCAACTGCTGTGATTGAA-----GATATATTGCCCGA 346
        |||
Db      101 ArgAlaThrAlaLeuIleSerLeuGlyLysAlaAlaIleGlyPheAspIlePheAlaArg 120
OY      347 GACAAATTCCTATTGTGTGGGAGGAGAAACAATTATACATTTGCAATCTGCTCGAAA 406
        |||
Db      121 AspLysIleProIleValIleGlyGlyThrAsnTyrTrpIleGlnSerLeuLeuTrpLys 140

```

QY 407 GTTCTGTGCAATACCAAGCCGAGAGATGGCACTGAGAAAGTATTGACCGAAAGTC 466
 Db 141 ValLeuValAsnThrIleGlyProGlnGlnMetGlyThrGlyValIleAspArgVal 160
 QY 467 GAGCTTGAAGAGAGAGAGCTGTACTTGCACAAAGCCCTAAGCCAGGTGACCCAGAA 526
 Db 161 GluLeuGlyGlyGlnAspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGln 180
 QY 527 ATGGCCCAAGCTCAGTCCATGACATGACAAACGCAAGAGTGGCAGAGCTTGCAAGTTT 586
 Db 181 MetAlaIalysLeuHisProHisAspLysArgLysValAlaArgSerLeuGlnAlaPhe 200
 QY 567 GAAAGAAACAGAAATCTCATAGTAATTTCTGCATGCGTCAACATACGGAAGAAGTGT 646
 Db 201 GlnGlyThrGlyIleSerHisSerIlePheLeuHisArgGlnHisThrGlnGlyGly 220
 QY 647 GGTCCCTGGAGGCTCCTGAAAGTTCTTACCCCTGACCTTGACCTTGCTGCTGCTGAC 706
 Db 221 GlyProLeuGlyGlyProLeuLysPheSerAsnProCysIleLeuThrLeuHisAlaAsp 240
 QY 707 CAGGAGTCTTGATGAGACGCTTGATGAAGAGGTGATGACATGCTTCTGCTGGGCTC 766
 Db 241 GlnAlaValLeuAspGlnArgLeuAspLysArgValAspAspMetLeuAlaGlyLeu 260
 QY 767 TTGGAGAACTAGAGATTTTCAACAGCCCTAATACAGAGAATGTTTGGAAATATAC 826
 Db 261 LeuGlnGlnLeuArgAspPheHisArgArgItyrAsnGlnLysAsnValSerGlnAsnSer 280
 QY 827 CAGAGCTATCAACATGATCTTCAATCAATGGCTTCAAGGAATTTTCAAGATACCTG 886
 Db 281 GlnAspArgGlnHisGlyIlePheGlnSerIleGlyPheLysGlnPheHisGlyItyrLeu 300
 QY 887 ATCACTGAGGAGAAATGCACTGAGAGACTAGTAAACAGCTTCTAAAGAAAGA----- 940
 Db 301 IleThrGlnGlyLysCysThrLeuGlnThrSerAsnGlnLeuLeuLysGlyIleGln 320
 QY 940 ----- 940
 Db 321 AlaLeuLysGlnValThrLysArgItyrAlaArgLysGlnAsnArgItyrValLysAsnArg 340
 QY 941 -----CCTGGCCCATGTTCCCTCCCTGCTATGAGGTATGCTGATGTC 988
 Db 341 PheLeuSerArgProGlyProIleValProProValItyrGlyLeuGlnValIleSerAspVal 360
 QY 989 TCGAAGTGGAGAGAGTCTGCTGTAACCTGCTTGAATGCTGCAAAATGTTCAATCAG 1048
 Db 361 SerLysTrpGlnGlnSerValLeuGlnProAlaLeuGlnIleValGlnSerPheIleGln 380
 QY 1049 GGCCCAAGCTCAGACCACTCCATTAAGATGCCATACATGAAGCTGAGAACAGAGA 1108
 Db 381 GlnHisLysProThrAlaThrProIleLysMetProItyrAsnGlnValGlnLysAsnArg 400
 QY 1109 AGTTATCACTGCTGACCTGCTGATGCAATCATCATGGGAGTCCGCAAGGGCAGG 1168
 Db 401 SerItyrHisLeuCysAspLeuLysAspArgIleIleGlyLysArgItyrGlnAlaAla 420
 QY 1169 CACATAAATCCCAATCCCACTTGAACCACTGAGAAAGAAAGAAAGATTTGACACTCAG 1228
 Db 421 HisIleLysSerLysSerHisLeuAsnGlnLeuLysArgArgArgLeuAspSerAsp 440
 QY 1229 GCTGTCAACACCATAGAAAGTCAAGAGTGTTCGCCAGACTATTAACAAAGACCTAAAGG 1288
 Db 441 AlaValAsnThrIleGlnLysGlnSerValIleSerProAspHisAsnLysGlnProLysGln 460
 QY 1289 AAGGATGCCCGAGGAGAAATGATCAAGCGTGAATGACAGCTT 1333
 Db 461 LysGlySerProGlyGlnAsnAspGlnGlnLeuLysCysSerVal 475

RESULT 3
 US-10-380-731-563
 : Sequence 563. Application US/10380731
 : GENERAL INFORMATION:

APPLICANT: Hyseq, Inc
 TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 FILE REFERENCE: 21272-114
 CURRENT APPLICATION NUMBER: US/10/380,731
 CURRENT FILING DATE: 2003-03-14
 PRIOR APPLICATION NUMBER: 09/659,671
 PRIOR FILING DATE: 2000-09-11
 NUMBER OF SEQ ID NOS: 888
 SOFTWARE: Custom
 SEQ ID NO: 563
 LENGTH: 411
 TYPE: PRP
 ORGANISM: Homo sapiens
 US-10-380-731-563

Alignment Scores:
 Pred. No.: 7,53e-153 Length: 411
 Score: 1989.00 Matches: 383
 Percent Similarity: 93.43% Conservative: 1
 Best Local Similarity: 93.19% Mismatches: 1
 Query Match: 55.64% Indels: 26
 DB: 6 Gaps: 1

US-09-513-151-3 (1-2041) x US-10-380-731-563 (1-411)

QY 179 ATGCAGGCTGTATGAAAGCCCTAGACATCATCCAAACAGCTTCTGCCAAGACAGAGA 238
 Db 1 MetGlnValItyrGlyGlyLeuAspIleIleThrAsnLysValSerIleGlnGlnArg 20
 QY 239 ATCTGCGGCGACCAATGATACAGCTTGGATGATCCCTGTGACCAATTAACAGAGG 298
 Db 21 IleCysArgHisHisMetIleSerPheValAspProLeuValThrAsnItyrThrValVal 40
 QY 299 GACTTCAGAAATAGAGCAACTGCTGATTTGAAGATATATTGTCGCGAGCAAAATTCCT 358
 Db 41 AspPheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIlePro 60
 QY 359 ATTGTGTGGAGAGAACCAATTAATTCATTTGAATCTCTGCTGGAAGTCTTGTCAAT 418
 Db 61 IleValIleGlyItyrHisAsnItyrIleGlnSerLeuLeuThrPlyValIleValAsn 80
 QY 419 ACCAAGCCCGAGAGATGGGCACTGAGAAAGTATGACGGAAGTGGAGCTTGAAG 478
 Db 81 ThrLysProGlnGlnMetItyrThrGlyLysValIleAspArgLysValIleLeuGlnLys 100
 QY 479 GAGAGTGTCTGTAATTCACAAACGCTAAGCCAGAGTGGACCCAGAAATGCTGCGCAAG 538
 Db 101 GluAspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaLys 120
 QY 539 CTGCATCAATGACAAACGCAAGAGTGGCCAGAGCTTGCAGAGTATTTGAAGAAACAGGA 598
 Db 121 LeuHisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGlnIleThrGly 140
 QY 599 ATCTGCAATGAAATTTCTCATGCTCAATACGTAAGAGAGTGGTGGTCCCTGGA 658
 Db 141 IleSerHisSerGlnPheLeuHisArgGlnHisThrGlnGlyGlyItyrProLeuGly 160
 QY 659 GGTCTCTGAAGTCTCTAACCCTTGACCTTGTGGCTTCAATGCTGACAGGACAGTTCTA 718
 Db 161 GlyProLeuLysPheSerAsnProCysIleLeuThrLeuHisAlaAspIleAlaValLeu 180
 QY 719 GATGAGCCTTGGATAGAGGCTGATGACATGCTTGTGCTGGGCTCTTGAAGAGACTA 778
 Db 181 AspGlnArgLeuAspLysArgValAspAspMetLeuAlaGlyLeuLeuGlnIleLeu 200
 QY 779 AAGAGATTTCAAGAGCCCTAATACAGAGAAATGTTTGGAAATTAACGAGAGACTATCA 838
 Db 201 ArgAspPheHisArgArgItyrAsnGlnLysAsnValSerGlnAsnSerGlnAspItyrGln 220
 QY 839 CATGATATCTTCAATCAATGCTGCTCAAGAGAAATTTACAGACTGATCTGATCAGAGA 898
 Db 221 HisGlyIlePheGlnSerIleGlyPheLysGlnPheHisGlyItyrLeuIleThrGlnGly 240


```

; APPLICANT: American Cyanamid Company, and Fletcher, Leah D., McMichael, John C.,
; APPLICANT: Russell, David P., and Zagursky, Robert J.
; TITLE OF INVENTION: Allostococcus ocellidis Open Reading Frames (ORFs) Encoding Polypep
; FILE REFERENCE: Application 1
; CURRENT APPLICATION NUMBER: PCT/US02/36123
; CURRENT FILING DATE: 2003-01-02
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4332
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Allostococcus ocellidis
PCT-0502-36123-4332

Alignment Scores:
Pred. No.: 2,296-23 Length: 315
Score: 393.00 Matches: 103
Percent Similarity: 55.59% Conservatve: 81
Best Local Similarity: 31.12% Mismatches: 113
Query Match: 10.99% Indels: 34
DB: 1 Gaps: 10

US-09-513-151-3 (1-2041) x PCT-US02-36123-4332 (1-315)
QY 83 CTTGATGATTCCTCGGGCCACGGGCGACCGCAATCCACGCTGGCGTTCGACTAGGC 142
DB 5 llellevalillevaliglyprothrghlyvaliglysthrlylseuuserleuglualea 24
QY 143 CAGCGGCTGGCGGCGTGCAGATCTCAGCGCTGCATCCAGGCTATGTAAGCCCTAGAC 202
DB 25 GlnAspleuAlaglyglillelleasnglyAspsermetgllyAllyrghlyleuasp 44
QY 203 ATCATCCACAAGAGTTTCTGCCCAAGACAGAGATCTGCCGGCACCATGATCAGC 262
DB 45 lleglythgilylvalthrghlysergllyAlaglylleprothshstleuasp 64
QY 263 TTGTGATCTCTGTGACCAATTCACAGAGTGCAGTGCATCCAGAAATAGACAATGCT 322
DB 65 lleleuglu---ValglnlnuAsprryservalserAspshneglnserlnghlyrmet 83
QY 323 CTGATTAAGATATATTTGCCGAGACAAATTCCTATTGTGTGGAGAAACCAATTAAT 382
DB 84 lysllellelallevalaleaserlysaAsnArgleuProillellevaliglylthrglyleu 103
QY 383 TACATTAATCTCTGCTCGAAGATCTT-----GTCAATACCAAGCCCAAGGAGAG 436
DB 104 TrrllegluserleuiletyrAspvalSerHAsglyLyAspAlaglnProAsprryglu 123
QY 437 GGCACCTGAGAAGATGATTCACCGAAGATGAGCTTGAAAGAGATGCTTGTACTT 496
DB 124 TrrAlaglyAlalyseuGlnAspleu-----AlAspelnlglnlyrghlyrlytle 140
QY 497 CACAAAGCCCTAAGCCAGGTGAGCCAGAAATGCTGCACAGCTGCATCCACATGACAA 556
DB 141 TrrglnleuLeuAsnArgleuAsprryAlaAlaAlalyllleHAsproAsnshnleu 160
QY 557 CGCAAGTGGCCAGAGCTTGCAGTCTTTTGAAGAAACGAGATCTCTATGTGATTT 616
DB 161 ValArgthrllleArgAlaleuGlnuValTrrHAsvalThrghlylthrhpheserSerphe 180
QY 617 CTCATTCGCAACATACGAGAGGTGGTGGTCCCTTGAGAGTTCCTGAGATTTCTCT 676
DB 181 -----GlnAspelnlysllysglyPro-----AlaleuTyr 191
QY 677 AACCTTCATCTTGGCTTCATGCTGCAGCAGAGATTCATGATGAGCGCTGGATAG 736
DB 192 GlnAlaphelellelleglyleuAsnThrAspArgserlyleuTrrAspArglleAsnThr 211
QY 737 AGGTGGATGACATCTTCTGCTGCGCTCTTGAGAGAACTAAGAGATTTTTCACAGAGCC 796
DB 212 ArgValAspGlnmetValAspLysglyleuGlnuGlnuVal-----Lys 226

```

```

QY 797 TATATCAGAGAGATTTTCGGAAATATACCGACAGATATCAACATGATATTCACATCA 856
DB 227 TrrpleuAlalyshthrryAsnAspAlaglnserArg-----Arggly 241
QY 857 ATTTGGCTTACAGAGATTTACACAGTACCTGATCAGTGCAGAAATGCACACTGGAGACT 916
DB 242 lleglyTrrArglnuValleuSerTrrleu-----AspelyAspelnThrphneglnu 259
QY 917 AGTAACAGCTCTTAAGAAAGAGAGCTGGTCCATGTCCCT-----GTCTATGGC 970
DB 260 AlalleArgAspillelyslslnsAsnserArgHsTrrAlalyAspGlnleuThrtrpHe 279
QY 971 TTAGAGTATCTGATCTCGAAGTGGAGAGACTGT-----CTTGAA 1015
DB 280 HlsAsnArgThrproValAspLysTrrTrryAspleuilleGlnAsprrygluGlnu 299
QY 1016 CCTGCTCTGAATCTGCGAAAGTTTCATCCAG 1048
DB 300 GlnuValillelyAspvallyAsnphleuGln 310

RESULT 9
US-09-950-084-4328
; Sequence 4328, Application US/09950084
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Robert S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/1C963052
; CURRENT APPLICATION NUMBER: US/09/950,084
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 09/417, 811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353, 718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266, 557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266, 556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266, 555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266, 542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266, 541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037, 934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036, 720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036, 338
; PRIOR FILING DATE: 1998-03-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 4328
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-950-084-4328

Alignment Scores:
Pred. No.: 5,526-19 Length: 319
Score: 339.00 Matches: 86
Percent Similarity: 55.64% Conservatve: 62
Best Local Similarity: 32.33% Mismatches: 92
Query Match: 9.48% Indels: 26
DB: 5 Gaps: 7

US-09-513-151-3 (1-2041) x US-09-950-084-4328 (1-319)
QY 83 CTTGATGATTCCTCGGGCCACGGGCGACCGCAATCCACGCTGGCGTTCGACTAGGC 142
DB 16 llellevalillevaliglyprothrghlyvaliglysthrlylseuuserleuglualea 35

```



```

QY 143 CAGCGCGTGGCGGTGAGATCTGACGGCTGACGAGGTATGAGGCGTAGAC 202
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36 LysArgIleAsnIleGluIleIleSerGlyAspSerMetClnValIleYrYrSHSmetAsn 55
QY 203 ATCATCACAACAAGTTCTGCGCCCAAGACAGAGAATCTGCGGACCCACATGATCAGC 262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56 IleGIYThrAlaLysValIleThrProGIuGluMetAspIleProIleSHSleuIleAsp 75
QY 263 TTGTGGATCCCTCTGTGACCAATTACACAGTGTGAGCTTCAGAAATAGACAACTGCT 322
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 IleLeuAsnProAspAspIlePheSerAlaTyroIleIleLysAspLeuAlaGluAsp 94
QY 323 CTGATTGAAGATATATTGGCCGAGACAAATCTCTATTGTGGGAGAGAACCAATTAT 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 95 LeuIleThrAspIleThrAsnArgGlyLysValProIleIleAlaGlyIleYrIleGlyLeu 114
QY 383 TACATTGAATCTCTGTCTGGAAAGTTCTTCATATACCAAGCCCGACGAGATGGGCACT 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 TyrIleGlnSerLeuIleTyraAsnTyroIleGluLysAspIleThrValIleProAlaGln 134
QY 443 GAGAAAGTGAATGACCGAAAGTGGAG---CTTGAAGAGAGAGATGCTTGAATCTCAC 499
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 135 LeuSerIleValIleLysGlnLysLeuSerAlaLeuGlnIleLeuAspAsnGlnIleLeu 154
QY 500 AAACGCGCTAAGCAGGTGAGCCGAAATGCTGCCAAGCTGCATCCACATGACAAACGC 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 AspTyroLeuAlaGlnPheAspAlaValSerAlaGluAsnIleIleSPROAsnAsnArgGln 174
QY 560 AAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGA-----ATCTCTCATATGGA 613
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 ArgValLeuArgAlaIleGluIleTyroIleLysThrIleLysLeuSerAsnArgLys 194
QY 614 TTTCCTCATGCTCAACATACAGGAAGAAGTGTGGTCCCTTGGAGTCTCTGAATTC 673
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 195 LysVal---GlnGlnPheThrIleGluAsn----- 202
QY 674 TCAACGCTTGCATCTTGGCTTCATGCTGACGAGCAGTTCATGAGCGCTTGAT 733
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 203 TyrAspIleLeuLeuLeuGlyIleGluMetSerArgLysIleYrSerArgIleAsn 722
QY 734 AAGAGGTGATGACATCTTGTCTGTGGCTTGTGGAGAACTAAGATTTTTCACAGA 793
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 LysArgValAspIleMetLeuAspIleGlyLeuPheArgGluVal----- 237
QY 794 CGCTATATCAGAAGATCTTTCGGAATAAGCCAGACATATCAACTGATATCTTCAA 853
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 238 -----GlnGlnLeuValIleGluGlnIleTyroIleIleSerGlnSer-----MetCln 252
QY 854 TCAATTGGCTTCAAGAA 871
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 AlaIleGlyTyrIleGlu 258

```

```

Score: 319.50 Matches: 88
Percent Similarity: 52.168 Conservative: 57
Best Local Similarity: 31.658 Mismatches: 92
Query Match: 8.948 Gaps: 41
DB: 6 Gaps: 11

US-09-513-151-3 (1-2041) x US-10-289-762-983 (1-342)
QY 83 CTGTAGTGAATCTGCGGGCCGAGCGGACCGGAAATCCAGCGTGGCGGTGACGCTAGGC 142
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34 IleValLeuLeuSerGlyProIleGlySerGlyIleThrAspValSerIleAlaLeuAla 53
QY 143 CAGCGCGTGGCGGTGAGATCTGACGGCTGACGAGGTATGAGGCGTAGAC 202
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 54 PrometIleAspIleGluIleValSerValAspSerMetClnValIleYrGlnIleMetAsp 73
QY 203 ATCATCACAACAAGTTCTGCGCCCAAGACAGAGAATCTGCGGACCCACATGATCAGC 262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74 IleGIYThrAlaLysValSerLeuLysAlaArgGlnIleIleProIleSHSleuIleAsp 93
QY 263 TTGTGGATCCCTCTGTGACCAATTACACAGTGTGAGCTTCAGAAATAGACAACTGCT 322
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 ---IleArgHisValIleGlnIleProIleAsnValAlaIleAspIleTyroIleAlaIleGln 112
QY 323 CTGATTGAAGATATATTGGCCGAGACAAATCTATTGTGGGAGAGAACCAATTAT 382
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 AlaCysGlnAsnIleLeuSerArgAsnLysValProIleLeuValIleGlySerGlyPhe 132
QY 383 TACATTGAATCTCTGTCTGGAAAGTTCTTCATATACCAAGCCCGACGAGATGGGCACT 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 TyrPheHisAlaPheLeu-----SerIleProIleGlyProIleAla 147
QY 443 GAGAAAGTGAATGACCGAAAGTGGAG---CTTGAAGAGAGAGATGCTT---GTACTT 496
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148 AspProGlnIleArgGlnIleGluIleGluAlaIleIleAlaGluIleValIleSerAlaLeu 167
QY 497 CACAACGCGCTAAGCAGGTGAGCCGAAATGCTGCCAAGCTGCATCCACATGACAA 556
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 168 TyrGlnAspLeuLeuLeuLysAspProGluIleAlaGlnIleThrIleThrLysAsnAspLys 187
QY 557 CGCAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGA-----ATCTCTCATATG 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 188 AsnLysIleIleArgGlyLeuGlnIleIleGlnIleLeuIleLysLysValSerAspHis 207
QY 611 GAATTT-----CTCCATGCTCAACATACGGAAGAAGTGGTGGCCCTT 655
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 GluTyrAspIleValIleProLysAlaSerArgIleTyroLys----- 220
QY 656 GGAAGTCTCTGAAGTTCTTAACCTTGGATGCACTTGG---CTTCATGCTGACCAAGCA 712
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 221 -----CysArgAlaIlePheLeuSerProGluThrGlu 231
QY 713 GTTCTAATGAGCGCTTGGATGAGAGGTGATGACATGCTGCTGCGGCTTGTGAG 772
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 232 PheLeuLysAsnAsnIleGlnMetArgCysGlnIleMetLeuGlnIleGlyLeuLeuGln 251
QY 773 GACTAAGAGATTTTCAACAGCGCTATATACAGAAGATGTTTGGAAAAATAGCCAGAGAC 832
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 252 GluValIleArgGlyLeu-----LeuAsnGln---GlyIleArgGluAsnPro----- 265
QY 833 TATCAACATGATCTTCCATCAATGCTTCAAGGAATTTCAAGAGTACCTG 886
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 266 -----SerAlaPheLysAlaIleGlyTyrArgIleGluIlePheLeu 280

```

Alignment Scores: 2.14e-17 Length: 342


```

OY 896 GGAATGACACCTGAGACCTAGTAACCGCTCTCAAGAA 937
DB 250 GylgInserLeuGlumetaIaValgluThrValIysGln 263

RESULT 13
US-09-134-000C-4383
; Sequence 4383, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-002
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4383
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4383

Alignment Scores:
Pred. No.: 3,26e-15 Length: 311
Score: 292.50 Matches: 84
Percent Similarity: 50.68% Conservative: 65
Best Local Similarity: 28.57% Mismatches: 102
Query Match: 8.18% Indels: 43
DB: 5 Gaps: 8

US-09-513-151-3 (1-2041) x US-09-134-000C-4383 (1-311)
OY 83 CTTGATGATTCCTGGGCGCCACGGGACCAATTCACGCTGCGCTTGCAGTAGGC 142
DB 4 ValIeuValIleValgluProthrAlaValgluIuAsnSerIleIysIrcysPheGcs 23
OY 143 CAGCGG--CTCGGGGATGATGCTGACGCTGCTCAGCTCAGCTATGAAGCCGA 199
DB 24 GlnIysIleLeuMetValIysLeuIleSerGlyAspSerMetGlnValIyrProIeuVal 43
OY 200 GACATCATCACCAAC--AAGCTTCTGCCCAAGACAGAGAAATCTCGCGGACCAATG 256
DB 44 ArgIeuThrProArgLeuIysSerThrGlyThrGlnIysGlnGlyPheProIleIysLeu 63
OY 257 ATGACCTTGTGATCTCTTGTGACCAATTCACAGTGTGCTGCTGAGAAATGAGCA 316
DB 64 IleAspCysGlnIuValSerGluThr--TyrSerAlaIleAspPheGlnIysGlnVal 82
OY 317 ACTGCTGATGATGAAGATATATTGCCGAGCAAAATTCATTTGTTGGAGAGACC 376
DB 83 ArgGlnIysIleIysGlnIleThrGlnIysGlnIysLeuProIleIleValgluIysThr 102
OY 377 AATTATTACATTGAATCTCTGCTGGAAGATTCTTGTCATATACCAAGCCCAAGAGATG 436
DB 103 GlyIeuThrIleGlnSerIleuThrIysAspPhe-----GlnIeu 115
OY 437 GGCATGAGAAAGTATGACCGA-----AAAGTGAGCTTGA 475
DB 116 GlySerArgGluIleAspSerProGluIleArgGluThrTyrAsnIeuPheAlaGlu 135
OY 476 AAGGAGATGCTCTTGTCTCACAACCGCTAAGCCAGGTGAGCCAGAAATGGCTGCC 535
DB 136 GluIysGlnIuValIleValIeuThrIleuIleGlnIysAspProIeuAlaIleAsn 155
OY 536 AAGCTGCATCCATGACAAACGAAAGTGGCCAGAGCTTGCACACTTTTGAAGAA 595
DB 156 SerIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 175
OY 596 GGAATCTGTCATAGTAATTTCTCATGTCATCAACATAGGAAGAGTGTGCTCCCT 655
DB 176 GlyIyrSerIleIeuThr-----ProIys 183

```

```

OY 656 GAGGCTCTCTGAACTCTCTAAACCTTGCACTTCCTTGGCTTCAAGTACAGGCACT 715
DB 184 GluIysProIleArgLeuIyrAspIyrTyrIleuIleGlnIuThrAspArgAlaIeu 203
OY 716 CTGATGAGCGCTGTGATGAAGAGGTGATGATGATCTGCTGCTGCTTGAAGAA 775
DB 204 LeuIyrGlnIuArgIleAsnGlnIyrValAspGlnMetMetThrGlnIuIleuIuGlu 223
OY 776 CTAAAGATTTTCACAGACGCTATATACAGAAATGTTTGGAAATATAGCAGCAT 835
DB 224 AlaIysGlnMetPheGlnIleProIleIleAlaIle----- 236
OY 836 CAACATGATATCTTCATCAATTCGCTTCAAGAAATTCACAGATACCTGATCAG 895
DB 237 -----GlnGlyIleIleIyrIyrGlnIuPheProIyrPhe-----Ser 249
OY 896 GGAATGACACCTGAGACCTAGTAACCGCTCTCAAGAA 937
DB 250 GylgInserLeuGlumetaIaValgluThrValIysGln 263

RESULT 14
US-10-156-761-10012
; Sequence 10012, Application US/10156761
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIDA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272657
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10012
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10012

Alignment Scores:
Pred. No.: 4.31e-15 Length: 312
Score: 291.00 Matches: 84
Percent Similarity: 50.53% Conservative: 58
Best Local Similarity: 29.89% Mismatches: 99
Query Match: 8.14% Indels: 40
DB: 6 Gaps: 9

US-09-513-151-3 (1-2041) x US-10-156-761-10012 (1-312)
OY 83 CTTGATGATTCCTGGGCGCCACGGGACCAATTCACGCTGCGCTTGCAGTAGGC 142
DB 10 ValIleIleValIeValgluProthrAlaIleGlnIysSerIleuIleuValIleIle 29
OY 143 CAGCGGCTGGGCGGTGAGATGCTGACGCTGATCATGCTGATTAAGCCCTAGAC 202
DB 30 GlnArgIleuAspGlyGluValIleAsnAlaAspSerMetGlnIleuIyrArgIleIleAsp 49
OY 203 ATCATCACCAACAGAGTTCTGCGCAAGAGAGATCTGCCGACCACTGATGAC 262
DB 50 IleGlyThrAlaIysLeuThrProGlnIuArgGlyIleProIleIleIleIleIle 69
OY 263 TTGTGATCTCTTGTGACCAATTAACAGTGTGAGACTTGAAGAAATGAGCAACTGCT 322
DB 70 IleThrAspValIleThrValIle--AlaSerValAlaGluIyrGlnIuArgIleuIleArgAla 88

```

```

QY 323 CTGATTGAGATATATTTCGCCGAGACAAATTCCTATTGTTGGGAGCAACATATT 382
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 89 ArgIILeaSPAlaLeuLeuAlaIleuLgIArgTTProIILeuValaIcIylSerGIleu 108
QY 383 TACATTGAATCTCTGCTGGAAGACTTCTTCAATACCAAGCCCGAGAGATGCGCACC 442
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 109 TyrValArgGLyAlaValaIaSPanLeu-----GluPheProGIlyThr 122
QY 443 GAGAAAGTATTCACCGAAAGTGCAGCTTGAAAGAGGATGCTT----- 490
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 123 ASPProGIlyVal-----ArgAlaArgLeuGIuGIuGIuLeuAlaLeuArgGIlyProGI 140
QY 491 GATCTTCACAAAGCGCTAGAGAGTGCAGCAAGTGCCTCCAGCAATGCATCCACT 550
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 141 AlaLeuHISAlaArgLeuAlaAlaAlaSPProGIlyAlaGIyHISAlaLeuSerSer 160
QY 551 GACAAAGCGAAAGTGCAGAGCTTGCAAGTTTGAAGAAACAGAAATCTCTCACTAGT 610
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 161 AaNGIlyArgArgIleValaArgAlaLeuGIuValaIleGIuIleThrGIy----- 176
QY 611 GAATTTCTCATGCTGCAACATACGGAAGAGTGTGCTCCCTTGAGAGTCCCTGAAG 670
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 177 -----ArgPhePheThrAlaAsn-----LeuProGIlyHISpSer 188
QY 671 TTCTTAACCCCTTGATCCTTGGCTTCATGCTGACAGCAAGTTAGATGAGCGCTTG 730
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 189 ValTyrAspThrValcIuIleGIyValaSPValaAlaArgProGIleuAspGIlyArgIle 208
QY 731 GATTAAGAGGTGATGATGATGCTTGTGCTGGGCTTGGAGAGAACTTAAGATTTTAC 790
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 209 AlaArgArgValaSPArgMetArgIuAlaGIyLeuValaSPolValaArg----- 225
QY 791 AGACGCTATATACAGAAGATGTTTGGCAAAATAGCCAGACTATCAACATGCT----- 844
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 226 -----AlaLeuGIuAlaGIuGIyLeuArgGIuGIyArgThr 237
QY 845 ANCTTCATCAATGCTTCAAGATTTCAAGACTGATCTGATCACTGAGCAAAATGC 904
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 238 AlaSerArgAlaLeuGIyArgGIuGIuValaLeuAlaLeu-----AlaGIyGIuGys 255
QY 905 ACA 907
    |||
DB 256 Thr 256

```

RESULT 15
PCT-US02-32727-18388
Sequence 18388, Application PC/TUS0232727

GENERAL INFORMATION:

APPLICANT: Mitcham, Jennifer

APPLICANT: Skelky, Yasir

APPLICANT: Persing, David

APPLICANT: Bhatia, Ajay

APPLICANT: Maisonneuve, Jean Francois

APPLICANT: Zhang, Yanni

APPLICANT: Wang, Siqing

APPLICANT: Jen, Shylan

APPLICANT: Lodes, Michael

APPLICANT: Benson, Darin

APPLICANT: Jones, Robert

APPLICANT: Carter, Brenda

APPLICANT: Douglas, John

TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes

FILE REFERENCE: 210121.51C1

CURRENT APPLICATION NUMBER: PCT/US02/32727

NUMBER OF SEQ ID NOS: 30992

SEQ ID NO 18388

LENGTH: 357

TYPE: PRT

ORGANISM: Propionl acnes

PCT-US02-32727-18388

Alignment Scores:

Pred. No.:	7,05e-15	Length:	357
Score:	288.50	Matches:	81
Percent Similarity:	50.89%	Conservative:	62
Best Local Similarity:	28.83%	Mismatch:	95
Query Match:	8.07%	Indels:	43
DB:	1	Gaps:	8

US-09-513-151-3 (1-2041) x PCT-US02-32727-18388 (1-357)

```

QY 77 CTACCTCTTGATGATTTCTCGGGCCACGGGCGCAAGCAATTCACGCTGGTGAG 136
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 47 LeuProValaIleCysLeuValcIyProThrAlaSerGIlySerGIyLeuAlaValArg 66
QY 137 CTAGCCACGGGCTC-----GGCGGTGATGCTGACGGCTACTGCTCATG 181
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 67 ValCysArgArgLeuTyrValaSPolHISProAlaGIuIleAlaSPnThrSPMet 86
QY 182 CAGGCTGTGAAGGCTGACATCATACCAACAAAGTTTCCGCCAGAGCAGAGAAATC 241
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 87 ValValTyrArgGIyMetAspIleGIyThrAlaThrProThrLeuArgGIuIleArgThr 106
QY 242 TCCCGCACACATGATGATGCTTGTGATCCTTGTGACCAATTCACAGTGTGAGC 301
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 107 ValValHISleuValSerIleLeuAspValThrValProSer---SerLeuValLeu 125
QY 302 TTCAGAAATAGCAATGCTGCTGATTAAGATATATTTGCCCGACAAATTCCTATT 361
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 126 MetGIuThrLeuAlaArgAspAlaValGIuAspCysLeuSerArgGIyValIleProVal 145
QY 362 GTTGGGAGGAGCAACATATATACATGATCTGCTGCAAGTCTTGATGATAC 421
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 146 LeuValcIyGIySerAlaLeuTyrThrLysAlaIleIleAspGIuMetSerIleProPro 165
QY 422 AAGCCCGAGAGATGGGCACTGAGAAAGTATTGACCGAAAGTGGAGCTTGAAGAGAG 481
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 166 ThrAspProGIlyValaArgAlaArg-----TyrGIuGIuGIyLeuAspAlaGIu 181
QY 482 GATGCTCTTGCTTCTACAAAGCCCTAAGCCAGGTGAGCCAGAAATGGCTGCCAAGCTG 541
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 182 GIyProArgValaLeuHISAspGIuLeuAlaArgAspProLysAlaIleGIuSerIle 201
QY 542 CATCCAGTACAAAGCGCAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGAGATC 601
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 202 LeuProGIlyAsnGIyArgArgIleValaArgAlaLeuGIuValaIleAspLeuThrGIySer 221
QY 602 TCTCATAGTGAATTT-----CTGCATGCTGCAACATACGGAAGAGGTGTGT 649
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 222 PheThrAlaThrIleProAspGIyThrLeuHISThrProLysThrValcIuMetGIy 240
QY 650 CCCCTTGAGAGTCCCTGTAAGTTCTGTACCTGATCCTTGGCTTCACTGCTAGCCAG 709
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 241 -----LeuGIuLeuSerArg 245
QY 710 GCAGTCTAGATAGAGCGCTTGAGATAAGAGGTGAGATGATGCTGCTGGCTGTG 769
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 246 LysAspIleAspGIuArgIleAlaSPArgValaSPGIuMetTyrAlaIleArgIlePheVal 265
QY 770 GAGGAACATAGAGATTTTCAAGAGCGCTTAATACAGAGAATGTTTGGAAAAATGACCG 829
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 266 AspGIuValaArgSerLeu-----AlaAsnValcIy 275
QY 830 GACTATCAACATGATCTTCCATCA-----ATTGGCTTCAAGGAATTTTCAAGATGAC 883
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 276 ---LeuArgGIuGIyLeuThrAlaSerArgAlaLeuGIyTyrArgGIuValaIleuGIuTyr 294
QY 884 CTG 886
    |||
DB 295 Leu 295

```

Search completed: April 21, 2003, 18:58:58
Job time: 99.4244 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 21, 2003, 18:51:16 : Search time 312.711 Seconds

(without alignments)
8416.087 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 3575

Sequence: 1 CTGCCCATAGATGCGTCGCGC.....TTTACACAGAAAAA 2041

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Dgapop 6.0	Dgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 9138288

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

```
-O=/cg2_1/USPTO.spool/US09513151/runat.15042003.141145.26426/app.query.fasta.1.2446
-DB=pending.patents_AA.Main -QFWT=fasten -SUPFIX=n2p.rapm -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09513151.ecgn.1.1.422.0runat.15042003.141145.26426
-NCPU=6 -ICPU=3 -NO_XMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=10 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
Pending_Patents_AA.Main:*
1: /cg2_6/ptodata/1/paa/US06.COMB.pep:*
2: /cg2_6/ptodata/1/paa/US06.COMB.pep:*
3: /cg2_6/ptodata/1/paa/US07.COMB.pep:*
4: /cg2_6/ptodata/1/paa/US08.COMB.pep:*
5: /cg2_6/ptodata/1/paa/US08.COMB.pep:*
6: /cg2_6/ptodata/1/paa/US08.COMB.pep:*
7: /cg2_6/ptodata/1/paa/US08.COMB.pep:*
8: /cg2_6/ptodata/1/paa/US08.COMB.pep:*
9: /cg2_6/ptodata/1/paa/US08.COMB.pep:*
10: /cg2_6/ptodata/1/paa/US08.COMB.pep:*
11: /cg2_6/ptodata/1/paa/US08.COMB.pep:*
12: /cg2_6/ptodata/1/paa/US08.COMB.pep:*
13: /cg2_6/ptodata/1/paa/US08.COMB.pep:*
14: /cg2_6/ptodata/1/paa/US09.COMB.pep:*
15: /cg2_6/ptodata/1/paa/US09.COMB.pep:*
16: /cg2_6/ptodata/1/paa/US09.COMB.pep:*
17: /cg2_6/ptodata/1/paa/US09.COMB.pep:*
18: /cg2_6/ptodata/1/paa/US09.COMB.pep:*
19: /cg2_6/ptodata/1/paa/US09.COMB.pep:*
20: /cg2_6/ptodata/1/paa/US09.COMB.pep:*
21: /cg2_6/ptodata/1/paa/US09.COMB.pep:*
22: /cg2_6/ptodata/1/paa/US09.COMB.pep:*
23: /cg2_6/ptodata/1/paa/US09.COMB.pep:*
24: /cg2_6/ptodata/1/paa/US10.COMB.pep:*
25: /cg2_6/ptodata/1/paa/US10.COMB.pep:*
26: /cg2_6/ptodata/1/paa/US10.COMB.pep:*
27: /cg2_6/ptodata/1/paa/US60.COMB.pep:*
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2253	63.0	467	1	PCR-US02-07826-153	Sequence 153, App
2	2253	63.0	467	24	PCR-US10-097-340-153	Sequence 153, App
3	2231	62.4	465	1	PCR-US02-07826-151	Sequence 151, App
4	2231	62.4	465	24	US-10-097-340-151	Sequence 151, App
5	1074	30.0	222	1	PCR-US01-01239-1394	Sequence 1394, App
6	1074	30.0	222	1	PCR-US01-01349-586	Sequence 586, App
7	1074	30.0	222	1	PCR-US01-01349-586	Sequence 586, App
8	1074	30.0	222	21	US-09-764-853-784	Sequence 784, App
9	1074	30.0	222	21	US-09-764-853-784	Sequence 784, App
10	1074	30.0	222	21	US-09-764-853-784	Sequence 784, App
11	1074	30.0	222	24	US-10-072-326-586	Sequence 586, App
12	1074	30.0	222	24	US-10-072-326-586	Sequence 586, App
13	686.5	19.2	143	21	US-09-757-028-2199	Sequence 2199, App
14	686.5	19.2	143	26	US-10-222-911-2199	Sequence 2199, App
15	597.5	16.7	252	16	US-09-270-767-44570	Sequence 44570, App
16	548	15.4	430	19	US-09-513-151-2	Sequence 2, App11
17	548	15.3	428	21	US-09-791-537-6482	Sequence 4482, App
18	536.5	15.0	427	21	US-09-791-537-6482	Sequence 4482, App
19	471	13.2	461	21	US-09-708-427-19098	Sequence 19098, App
20	456	12.8	451	21	US-09-708-427-19098	Sequence 19098, App
21	403.5	11.3	314	21	US-09-791-537-70445	Sequence 70445, App
22	391	10.9	423	21	US-09-708-427-19098	Sequence 19100, App
23	389	10.9	306	15	US-09-107-532A-5131	Sequence 5131, App
24	389	10.9	306	15	US-09-107-532A-5131	Sequence 5131, App
25	386.5	10.8	386	16	US-09-248-796-17815	Sequence 17815, App
26	386.5	10.8	386	27	US-60-096-409-17815	Sequence 17815, App
27	367	10.3	257	25	US-10-155-881-8448	Sequence 8448, App
28	355	9.9	329	21	US-09-710-279-1460	Sequence 1460, App
29	339	9.5	319	12	US-08-827-326-288	Sequence 3288, App
30	339	9.5	319	20	US-09-611-529-4328	Sequence 4328, App
31	338	9.5	312	21	US-09-708-427-17588	Sequence 27588, App
32	338	9.5	318	21	US-09-708-427-17588	Sequence 27587, App
33	334.5	9.4	311	17	US-09-368-382-32	Sequence 32, App1
34	334.5	9.4	311	20	US-09-649-145-32	Sequence 32, App1
35	334.5	9.4	311	26	US-10-260-877-32	Sequence 32, App1
36	333.5	9.3	329	19	US-09-513-966A-62990	Sequence 62990, App
37	327.5	9.2	290	19	US-09-513-966A-62991	Sequence 62991, App
38	324	9.1	308	21	US-09-708-427-27589	Sequence 27589, App
39	320	9.0	576775	12	US-08-895-611-2	Sequence 2, App11
40	320	9.0	576775	12	US-08-895-611D-2	Sequence 2, App11
41	320	9.0	576775	18	US-09-458-180-2	Sequence 2, App11
42	320	9.0	576775	22	US-09-895-611D-2	Sequence 2, App11
43	319.5	8.9	329	19	US-09-513-966A-18307	Sequence 18307, App
44	319.5	8.9	342	15	US-09-198-452A-983	Sequence 983, App
45	319.5	8.9	345	18	US-09-438-185-912	Sequence 912, App

ALIGNMENTS

```
RESULT 1
PCT-US02-07826-153
; Sequence 153, Application PC/US0207826
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030PC
; CURRENT APPLICATION NUMBER: PCT/US02/07826
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/276,026
```

Db 221 PheSerAsnProCysIleLeuTrpLeuHisAlaAspGlnAlaValLeuAspLysIleArgLeu 240

Qy 731 GATAAGAGGGGTGGATGACATGCTGCTGCTGGGGCTCTTGAGAGCACTAAGAAATTTTCAc 790

Db 241 AspLysArgValAspAspMetLeuAlaIleGlyLeuLeuGlnIleuArgAspPheHis 260

Qy 791 AGACGCTTAATGCACAAACAATGTTCTCGAAATATGACAGAGACTATCAACATGGTAATCTTC 850

Db 261 ArgArgLysAsnGlnLysAsnValSerGlnAsnSerGlnAspTyrGlnHisGlyIlePhe 280

Qy 851 CAATCAATTTGGCTTCAGAAATTTTCACGAGTACCTGATCACTCACTGAGGAGAAATACACACTG 910

Db 281 GlnSerIleGlyPheLysGlnPheHisGlnIleuArgLeuIlePheGlnGlyLysCysThrLeu 300

Qy 911 GAGACTAGTAAACACAGCTTCTTAAGAAGA----- 940

Db 301 GlnThrSerAsnGlnLeuLeuLysLysGlyIleGlnAlaLeuLysGlnValThrLysArg 320

Qy 941 -----CCTGGTCCCACT 952

Db 321 TyrAlaArgLysGlnAsnArgTyrValLysAsnArgPheLeuSerArgProGlyProIle 340

Qy 953 GTCCCCCTGTCTATAGGCTTAGAGGTAATGATGTCTCGAACTGGGAGAGACTGTCTTT 1012

Db 341 ValProProValTyrGlyLeuGlnValSerAspValSerLysTyrProLysLeuValLeu 360

Qy 1013 GAACCTGCTCTTGAATGCTGCAAAAGTTTCATCCAGGGCCACAGGCTTACAGCCACTCCA 1072

Db 361 GluProAlaLeuGlnIleValGlnSerPheIleGlnGlnLysPheProThrAlaThrPro 380

Qy 1073 ATAAAGATGCCATACAAATGAGCTGAGAACAAAGAGAGTTATCACCTGTGTGACCTCTGT 1132

Db 381 IleLysMetProTyrAsnGlnIleGlnLysAsnLysArgSerTyrHisLeuLysAspLysCys 400

Qy 1133 GATCAATCATCATGTTGGGGATGCGCAATGGGAGGCGACATAAATCCAAATCCCACTTG 1192

Db 401 AspArgIleIleIleIleGlyAspArgIleuTrpAlaIleHisIleLysSerLysSerHisLys 420

Qy 1193 AACCAACTGAGAAAGAAGAAGATTGGACTCAGATGCTGTCAACACSCATAGAAAGTCAg 1252

Db 421 AsnGlnLeuLysLysArgIleArgIleAsnAspSerAspAlaValAsnThrIleGlnSerGln 440

Qy 1253 AGTGTTCCTCCAGACTATATACAAAGACCTAAAGGAGAGGAGTCCCGAGGCGAATGAT 1312

Db 441 SerValSerProAspHisAsnLysGlnProLysGlnLysGlySerProGlyGlnAsnAsp 460

Qy 1313 CAAGAGCTGAATGACAGCGTT 1333

Db 461 GlnGlnLeuLysCysSerVal 467

RESULT 2
US-10-097-340-153
Sequence 153: Application US/10097340
GENERAL INFORMATION:
APPLICANT: John MONAHAN
APPLICANT: Manjula GANNANVARAPU
APPLICANT: Sebastian HOESCH
APPLICANT: Shubhangi KAMATKAR
APPLICANT: Steve G KOVATS
APPLICANT: Rachel E. MEYERS
APPLICANT: Michael MORRISEY
APPLICANT: Peter OLANDT
APPLICANT: Ami SEN
APPLICANT: Peter VEIBY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. BAST, JR.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHMANDT
APPLICANT: Xunlei ZHAO
APPLICANT: Karen GIANT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification
OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
FILE REFERENCE: MRI-030

```

CURRENT APPLICATION NUMBER: US/10/097.340
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 153
LENGTH: 467
TYPE: PRT
ORGANISM: Homo sapiens
US-10-097-340-153

Alignment Scores:
Pred. No.: 1,03e-210 Length: 467
Score: 2253.00 Matches: 439
Percent Similarity: 94.22% Conservative: 1
Best Local Similarity: 94.00% Mismatches: 1
Query Match: 63.02% Indels: 26
DB: 24 Gaps: 1

US-09-513-151-3 (1-2041) x US-10-097-340-153 (1-467)
OY 11 AAGCGTCGCGGCTGCTGACAGAGATTCCTGCGAGTGGCGCTGAGGCGCTGCAA 70
DB 1 MetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArgGlyLeuGln 20
OY 71 CGAGCCCTACCTCTTGTAGTATTCYGGGGCCAGCGGCGACCGGCAATCCAGCTGGCG 130
DB 21 ArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSerThrLeuAla 40
OY 131 TTGCAGTAAAGCCAGCGGCTGGCGGTGAGATGTCAGCGCTACTGCATGCAAGTCTAT 190
DB 41 LeuGlnLeuGlyGlnArgLeuGlyGlyGlnIleValSerAlaAspSerMetGlnValTyr 60
OY 191 GAAGGCTAGACATCATCAACCAAGGTTCTGCCCAAGAGAGAGAACTCCGCGCAC 250
DB 61 GluGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnIleArgIleCysArgHis 80
OY 251 CACATATCAGCTTGTGATCCTCTGTGACCAATTAACAGTGTGAGTTCAGAAAT 310
DB 81 HisMetIleSerThrValAspProLeuValThrAsnTyrThrValValAspPheArgAsn 100
OY 311 AGAGCAAGCTGCTGATGAAGATATATTTGCCGAGACAAATTCCTATTTGTGGGA 370
DB 101 ArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIleValIleGly 120
OY 371 GGAACAATTAATACATTAATCTGCTGCTGGAAGTCTTGTGCATATCAAGCCGAG 430
DB 121 GATTTAAsnTyrTyrIleGlySerLeuLeuTyrLysValIleuValAsnThrLysProGln 140
OY 431 GAGATGGGCACTGAGAAAGATGATGACCAAAAGTGGAGCTGAAAAGAGAGATGCTT 490
DB 141 GluMetGlyThrGlnLysValIleAspArgLysValIleuGlnLysGlnAspGlyLeu 160
OY 491 GTAATTCACAACGCTTAAGCAGGTGACCCAGAAATGCTGCCAAGTGCATGCAT 550
DB 161 ValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaIleLysLeuHisProHis 180
OY 551 GAAACGCAAAAGTGGCCAGAGCTTCCAGATTTTGAAGAAACAGAAATCTCCATAGT 610
DB 181 AspLysArgLysValAlaArgSerLeuGlnValPheGlnGlnIleThrGlyIleSerHisSer 200

```

```

OY 611 GAATTTCTCCATGCTCAACATACGGAAGAAGTGGTGGTCCCTTGAGGCTCTGAG 670
DB 201 GluPheLeuHisArgGlnHisThrGlnGlnGlyGlyProLeuGlyLysProLeuLys 220
OY 671 TTCTCTAACCTTGCATCTTGGCTCTGCTGAGGAGGAGGAGTTCATATGAGCGCTTG 730
DB 221 PheSerAsnProCysIleLeuThrPheHisAlaAspGlnAlaValLeuAspGlnArgLeu 240
OY 731 GATAAGAGGTGATGACATGCTTGGCTGCGCTTGTGAGGAACTAAGATTTTCAC 790
DB 241 AspLysArgValAspSerPheLeuAlaIleGlyLeuLeuGlnClnIleuArgAspPheHis 260
OY 791 AGAGCTATATATCAGAAGATGTTGGGAAATAGCCAGACATATCAATGATCTTC 850
DB 261 ArgArgTyrAsnGlnLysAsnValSerGlnAsnSerGlnAspTyrGlnHisGlyIlePhe 280
OY 851 CAACTAATGGCTTCAAGGAATTCACGAGTACCTGATCAGTGAAGGAAATGACACTG 910
DB 281 GlnSerIleGlyPheLysGlnPheHisGlyTyrLeuIleThrGlnClnLysCysThrLeu 300
OY 911 GAGACTAGTAACCGCTTCAAGAAAGAA-----CTGCTCCCAT 940
DB 301 GluThrSerAsnGlnLeuLysGlyIleGlnAlaLeuLysGlnValThrLysArg 320
OY 941 -----CTGCTCCCAT 952
DB 321 TyrAlaArgLysGlnAsnArgTyrValLysAsnArgPheLeuSerArgProGlyProIle 340
OY 953 GTCCCTCCCTGATGCTGATGAGTATGATGATGATGATGATGATGATGATGATGATGAT 1012
DB 341 ValProValTyrGlyLeuGlnValSerAspValSerLysThrGlnGlnSerValLeu 360
OY 1013 GAACCTGCTCTGAAATCGTCAAGTTCATCCAGGCGCCACAGCTTACAGCCACTGCA 1072
DB 361 GluProAlaLeuGlnIleValGlnSerPheIleGlnGlyHisLysProThrAlaThrPro 380
OY 1073 ATAAGATGCCATPACATGAGCTGAGACACAGAAATATACCTGCTGACCTCTGT 1132
DB 381 IleLysMetProTyrAsnGlnIleGlnAsnLysArgSerTyrHisLysCysAspLeuCys 400
OY 1133 GATCGAATCATCAFTGGGAGTTCGGAATGGGCGGCGACATATAATCCAAATCCACTTG 1192
DB 401 AspArgIleIleIleGlyAspArgGlnTyrAlaAlaHisIleLysSerLysSerHisLeu 420
OY 1193 AACCACTGAAGAAAGAAAGAAATGAGTCAAGTCTGCAACACATAGAAAGTCAAG 1252
DB 421 AsnGlnLeuLysLysArgArgArgLeuAspSerAspAlaValAsnThrIleGlnSerGln 440
OY 1253 AGTGTTCCTCCAGACTATTAACAAAGAACTTAAGGGAAGGATCCGAGGCGAAGAT 1312
DB 441 SerValSerProAspHisAsnLysGlnProLysGlnLysGlySerProGlnGlnAsnAsp 460
OY 1313 CAAGAGCTGAATGCAAGCTGTT 1333
DB 461 GlnGlnLeuLysCysSerVal 467

RESULT 3
PCT-US02-07826-151
Sequence 151, Application PC/TUS0207826
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc. et al.
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
FILE REFERENCE: MRI-030PC
CURRENT APPLICATION NUMBER: PCT/US02/07826
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14

```



```

: PRIOR APPLICATION NUMBER: 60/324,967
: PRIOR FILING DATE: 2001/09/26
: PRIOR APPLICATION NUMBER: 60/311,732
: PRIOR FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: 60/325,102
: PRIOR FILING DATE: 2001-09-26
: PRIOR APPLICATION NUMBER: 60/323,580
: PRIOR FILING DATE: 2001-09-19
: NUMBER OF SEQ ID NOS: 363
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 151
: LENGTH: 465
: TYPE: PRT
: ORGANISM: Homo sapiens
PCT-US02-07826-151

Alignment Scores:
Pred. No.: 1,47e-208      Length: 465
Score: 2231.00           Matches: 437
Percent Similarity: 93.79%    Conservative: 1
Best Local Similarity: 93.58%   Mismatches: 1
Query Match: 62.41%          Indels: 28
                                Gaps: 2

US-09-513-151-3 (1-2041) x PCT-US02-07826-151 (1-465)

QY 11 ATGGCGTCGCGGCCGTGCACAGCAAGTTCCTTGCGCAGTAGCGGCTCAGGGCCCTCAA 70
Db 1 MetAlaserValAlAlAlAlAlArgAlAlProValIdlySerGIyleuAArgIyleuGIln 20

QY 71 CGGACCCTACTCTTTGAATGATATTCTCGGGCCACGGGGCACCGGCAATCCAGCTGGCG 130
Db 21 ArgThrLeuProLeuValAlAlIleLeuGIyAlAthrGIlyThr-----SerThrLeuAla 38

QY 131 TTGAGCTAGGCGCCAGCGGCTCGCGGGAGAGATGCTGACGGCTGACTCCATGACAGCTAT 150
Db 39 LeuGIlnLeuGIlnArgLeuGIylGIyGIlnLeuAlaISerAlAspSermetGIlnAlTyrr 58

QY 191 GAAGCGCTAGCATCATCACCAACAAGGTTTCTGCCAAGAGCAGAAATCTGCCGCGCAC 250
Db 59 GluGIyleuAspIleIleThrAsnLysValSerAladInGIlnArgIIecysArgHis 78

QY 251 CACATGATCAGCTTTGTGGATCCTCTTGTGACCAATTACACAGTGTGACTTCAGAAT 310
Db 79 HisMetIleSerPheValAspProLeuValIThrsnITyrThrValAlaIsppheArgsn 98

QY 311 AGACAACACTGCTCGATTGGAAGATAATATTGGCCGACACAANAATTCCTATTGTTGGGA 370
Db 99 ArgAlathrAlaLeuIleGIlnAspIlePhehelArgrAspsIleProIleValAlGIly 118

QY 371 GGAACCAATTAATTACATTTGAATCTCTGCTGTGAAAATTTCTTGCAATACCAAGCCCAG 430
Db 119 GlyIhrAsnITryITrIleGIlnSerLeuIleITrpIlystAlLeuValAsnIThrIlyspIoGIln 138

QY 431 GAGATGGGCACGTGAAAAGTGATTTGACCGAAAAAGTGGAGGTTGAAAAGAGATGCTTT 490
Db 139 GIlnmetGIlyIThrGIlnLysValIleAspArgLyseAlGIlnLeuGIlnLysGIlnAspGIlyLeu 158

QY 491 GTACTTCACAAAGCCCTAACGSCAGAGTGGACCCGAAATGGCTGCTCCAAGCTGCATCCACAT 550
Db 159 ValLeuIhnIstAsArgLysEuseGIlnValAspProGIlnmetAlAlAlLysLysIhnAspRoHis 178

QY 551 GACAACAAGCAAGTGGCCAGAGCTTGCAAGTTTGTGAAAAGAACAGAAATCTCTCAATAGT 610
Db 179 AspIysArgrLysValAlAlatArgsetLeuGIlnValPheIlnIdIThrGrIlyIleSerHisSer 198

QY 611 GAATTTTCTCATGCTGCAACAATPAGGAAAGAAAGTGGTGGTCCCTTGAGAGTCTCTGAAG 670
Db 199 GIlnPheLeuIhnAsrArgGIlnIstHrGIlnGIylGIyGIyProLeuGIylGIyProLeuLys 218

QY 671 TTCTCTAAACCTTGCAATCTTTGGCTTCAGATGCTGACAGGCATTTGATATGAGCGCTTG 730
Db 219 PheSerAsnProCysIleLeuIThrPleuIhnIstAlAspRGIlnAlAlValIleuAspRGIlnArgLeu 238

```

```

OY 731 GATTAAGAGGCTGATACATACATGCTGCTGCTGGGCTTTGGAGCAACTAAGATTTCAC 790
Db 239 ASPLYSATGVALASPRSPMCTLEUALAALAGLYLENDLCLNGLULEULATGASPRPHENS 258
OY 791 AGACGCTATAATCAGAAAGATGTTTCGCAAAATATGCGCAGACTATCAATGCTATCTTC 850
Db 259 ATGATGTYTASNGLINDYSANVALSERGLUANSERGLNASPRYRINDINHSGLIEPHE 278
OY 851 CAATCAATTGCGTTCAAGGAATTTTCAGAGTACCTGATCACTGAGGAAATGACACSTG 910
Db 279 GINSERTIEGLYPHELYSGILUPHENISGLUTYRLEUETITRGLNGLYSCYSTHLEU 298
OY 911 GAGACTAGTAACACGCTTCTAAAGAAAGAA----- 940
Db 299 GUTTHRSERASNGLINDLEULSYLSGLYILEGLUALALEULSGLINVALTHRLYSARG 318
OY 941 -----CCTGGTCCACTT 952
Db 319 TYRILATARGLYSGLINASNAITGRVALLYYSANATGPHLEUSENATGRPROGLPROLIE 338
OY 953 GTCCCCCTGTCTATAGGCTTTAAGGTAATGTGATGTCGCAAGTGGGAGGAGTCTGTCTT 1012
Db 339 VALPROFOVALTYGLYGLYENGLUVALSERASVALSERLYSTRPLUGLUSEVALLEU 358
OY 1013 GAACCTGCGCTTGAATTCGTGCAAGTTTCATCCSAGGCGCACAGGCTTACAGCACTCCA 1072
Db 359 GUPTROALALENGLILLEVALINSERPHETIEGLINGLYNHILYSPROTTHRALATPRPO 378
OY 1073 ATAAGATGCGCTTACATATAGCTGAGACAGAGAGAGTATACCTGTGTGACCTGTGT 1132
Db 379 ILELYSMERPTOTYRASNGLINDLALUANANLYSTRGSETTYNHISLEUCYSAPRLEUCYS 398
OY 1133 GATGCAATCATCATTTGGGGATGCGGANTGGCAGCGCACATAAATCCAAATCCCACTTG 1192
Db 399 ASPATGILELLELLEGLYSPARGIUTRPAALALANLILLEYSERLYSSERHISLEU 418
OY 1193 AACCAACTGAAAGAAAGAAAGATTGAGCTGAGATGCTGCTCAACCCATTAAGAACTGAG 1252
Db 419 ASNGLINDLEULYSLYSATGTGTATGTAUASPSERASPRALAVALASNTHRIIEGLUSERGIL 438
OY 1253 AGTGTTCCTCCAGATATATACCAAGAACCTTAAGGAGAGGATCCCCAGGCGAGAAATGAT 1312
Db 439 SERVALSERPROASPHISANLSNLSGLINUPROLYSGILUGLYSERPROLYGLINASNAP 458
OY 1313 CAAGAGCTGAATCGACGCTT 1333
Db 459 GINGLULEULYSCYSEVAL 465

RESULT 4
US-10-097-340-151
; Sequence 151, Application US/10097340
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNANARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangt KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340

```



```

; COMMENT FILING DATE: 2001-01-14
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939

```

```

; Sequence /84, Application PC/T0S0101349
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.

```

FILE REFERENCE: PUZ06PCT
CURRENT APPLICATION NUMBER: PCT/US01/01349
CURRENT FILING DATE: 2001-01-14
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 939
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 784
LENGTH: 222
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (124)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (145)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01349-784

Alignment Scores:
Pred. No.: 2,06e-95 Length: 222
Score: 1074.00 Matches: 214
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.17% Mismatches: 3
Query Match: 30.04% Indels: 1
Gaps: 0
DB:

US-09-513-151-3 (1-2041) x PCT-US01-01349-784 (1-222)

OY 2 TGCCATAAGATGCGTCGCGGCGCTGCACAGCAGTTCCTGTGGCAGTGGGCTCAGG 61
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
OY 62 GGCTGACAGGACCCCTACCTTGTATGATTCCTGGGGCCACGGGACCGGCAAAATCC 121
Db 21 GlyLeuGlnArgThrLeuProLeuValAlaIleLeuGlyAlaThrGlyThrGlyLysSer 40
OY 122 ACGCTGGCGTTGCACCTAGCGCCAGCGGCTGGAGATGTCAGCGCTACCTCATTG 181
Db 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGlyIleValSerAlaAspSerMet 60
OY 182 CAGGTCTATGAAGGCTAGACATCATCAACAAGGTTTCTGCCCAAGACGAGAATC 241
Db 61 GlnValTyrGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80
OY 242 TGGCGGACACCATGATCAGCTTTGTGATCCTCTTGACCAATTACAGTGGTGAC 301
Db 81 CysArgHisHisMetCileSerPheValAspProLeuValThrAsnTyrThrValAlaAsp 100
OY 302 TTCGAAATAGCAACAGCTCTGATGAAGATATATTGCCCGAGCAAAATTCCTATT 361
Db 101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120
OY 362 GTTGTGGAGAGAACCAATATTATACATGATCTCTCTGTGAAAGTTCTTGCAATAC 421
Db 121 ValValIgly***ThrAsnTyrTyrIleGlnSerLeuLeuTrrplyValLeuValAsnThr 140
OY 422 AAGCCCCAGAGATGGGACCTGAGAAAGTATGACCCGAAAAGTGGAGCTTGAAGAAG 481
Db 141 LysProGlnGlu***GlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysGln 160
OY 482 GATGGCTTTGACTTCACAAACGGCTAGCGGCGAGGCCGGAAGTGGCTGCACAGCG 541
Db 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetCalaIaLysLeu 180
OY 542 CATCCACATGACAAACGGAAGTGGCCAGAGCTTGCAAGTTTGTGAAGAAACAGGAATC 601
Db 181 HisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGlnGlnIleThrGlyIle 200
OY 602 TCTCATAGTGAATTTCTCCATCCTGACATAC- GGAAGAAGGTGGTGGTCCCT 654
Db 201 SerHisSerGlnPheLeuHisArgGlnHisThrGlyArgLysTrrPrrSerPro 218

RESULT 8
US-09-764-853-586
Sequence 586, Application US/09764853
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PUZ06
CURRENT APPLICATION NUMBER: US/09/764,853
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 939
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 586
LENGTH: 222
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (124)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (145)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-586

Alignment Scores:
Pred. No.: 2,06e-95 Length: 222
Score: 1074.00 Matches: 214
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.17% Mismatches: 3
Query Match: 30.04% Indels: 1
Gaps: 0
DB:

US-09-513-151-3 (1-2041) x US-09-764-853-586 (1-222)

OY 2 TGCCATAAGATGCGTCGCGGCGCTGCACAGCAGTTCCTGTGGCAGTGGGCTCAGG 61
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
OY 62 GGCTGACAGGACCCCTACCTTGTATGATTCCTGGGGCCACGGGACCGGCAAAATCC 121
Db 21 GlyLeuGlnArgThrLeuProLeuValAlaIleLeuGlyAlaThrGlyThrGlyLysSer 40
OY 122 ACGCTGGCGTTGCACCTAGCGCCAGCGGCTGGAGATGTCAGCGCTACCTCATTG 181
Db 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGlyIleValSerAlaAspSerMet 60
OY 182 CAGGTCTATGAAGGCTAGACATCATCAACAAGGTTTCTGCCCAAGACGAGAATC 241
Db 61 GlnValTyrGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80
OY 242 TGGCGGACACCATGATCAGCTTTGTGATCCTCTTGACCAATTACAGTGGTGAC 301
Db 81 CysArgHisHisMetCileSerPheValAspProLeuValThrAsnTyrThrValAlaAsp 100
OY 302 TTCGAAATAGCAACAGCTCTGATGAAGATATATTGCCCGAGCAAAATTCCTATT 361
Db 101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120
OY 362 GTTGTGGAGAGAACCAATATTATACATGATCTCTCTGTGAAAGTTCTTGCAATAC 421
Db 121 ValValIgly***ThrAsnTyrTyrIleGlnSerLeuLeuTrrplyValLeuValAsnThr 140
OY 422 AAGCCCCAGAGATGGGACCTGAGAAAGTATGACCCGAAAAGTGGAGCTTGAAGAAG 481
Db 141 LysProGlnGlu***GlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysGln 160
OY 482 GATGGCTTTGACTTCACAAACGGCTAGCGGCGAGGCCGGAAGTGGCTGCACAGCG 541
Db 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetCalaIaLysLeu 180
OY 542 CATCCACATGACAAACGGAAGTGGCCAGAGCTTGCAAGTTTGTGAAGAAACAGGAATC 601

```

Db      181 HisProHisAspLysArgValAlaArgSerLeuGlnValPheGluIuThrGlyIle 200
QY      602 TCTCATAGTGAATTTCTCATGTCACACATAC-GGAGAAGGTGGTCCCT 654
Db      201 SerHisSerGluPheLeuHisArgGlnHisThrGlyArgLysTrpIlePro 218

RESULT 9
US-09-764-853-784
; Sequence 784, Application US/09764853
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1206
; CURRENT APPLICATION NUMBER: US/09/764,853
; PRIOR APPLICATION DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 784
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-784

```

```

Alignment Scores:
Pred. No.: 2,06e-95 Length: 222
Score: 1074.00 Matches: 214
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.17% Mismatches: 3
Query Match: 30.04% Indels: 1
DB: 21 Gaps: 0

```

US-09-513-151-3 (1-2041) x US-09-764-853-784 (1-222)

```

QY      2 TGCATTAAGATGCGCTCCGCGGCTGCACGAGCATGTCGTGGGCGAGTGGCTCAGG 61
Db      1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
QY      62 GGCCTGCAAGGAGACCCCTACTCTTGATGATTCCTGGGGCCAGCGGCAATCC 121
Db      21 GlyLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40
QY      122 AGCGTGGCGCTTCAGCTAGACCGGCGGCGGTGAGATGCTCAGCGCTACTCCATG 181
Db      41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlnIleValSerIleAspSerMet 60
QY      182 CAGGTCTATGAAGGCTTACATCATCATCACCAAGGTTCTGCCAAGAGAGAAATCC 241
Db      61 GlnValIlyrGluGlyLeuAspIleIleThrAsnLysValSerAlaGlnIuGlnArgIle 80
QY      242 TGCCGGCACCATATGATCAGCTTGTGATCCTCTTGACCAATTAACAGTGGTGAC 301
Db      81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnIlyrThrValValAsp 100
QY      302 TTCAGAAATAGCAACTGCTGTGATGAAGATATATTGCCCGAGACAAATTCCTATT 361
Db      101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120
QY      362 GTTGTGGGAGAGAACCAATATTATACATGATCTGCTGTGACCAATTAACAGTGGTAC 421
Db      121 ValValGly***ThrAsnIlyrIleGlnSerLeuLeuTrpLysValLeuValAsnThr 140
QY      422 AAGCCCCAGAGATGGGCACTGAGAAGATGATGACGAAAGTGGAGCTTGAAGAGAG 481
Db      141 LysProGlnGlu***GlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysGln 160

```

```

QY      482 GATGGCTTGTACTTTCACAAAGCCTAAGCCAGGTGAGCCAGAAATGGCTGCCAAGTG 541
Db      161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaLysLeu 180
QY      542 CATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGAAATC 601
Db      181 HisProHisAspLysArgValAlaArgSerLeuGlnValPheGluIuThrGlyIle 200
QY      602 TCTCATAGTGAATTTCTCATGTCACACATAC-GGAGAAGGTGGTCCCT 654
Db      201 SerHisSerGluPheLeuHisArgGlnHisThrGlyArgLysTrpIlePro 218

RESULT 10
US-09-764-902-1394
; Sequence 1394, Application US/09764902
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1213
; CURRENT APPLICATION NUMBER: US/09/764,902
; PRIOR APPLICATION DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 2318
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1394
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-902-1394

```

```

Alignment Scores:
Pred. No.: 2,06e-95 Length: 222
Score: 1074.00 Matches: 214
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.17% Mismatches: 3
Query Match: 30.04% Indels: 1
DB: 21 Gaps: 0

```

US-09-513-151-3 (1-2041) x US-09-764-902-1394 (1-222)

```

QY      2 TGCATTAAGATGCGCTCCGCGGCTGCACGAGCATGTCGTGGGCGAGTGGCTCAGG 61
Db      1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
QY      62 GGCCTGCAAGGAGACCCCTACTCTTGATGATTCCTGGGGCCAGCGGCAATCC 121
Db      21 GlyLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40
QY      122 AGCGTGGCGCTTCAGCTAGACCGGCGGCGGTGAGATGCTCAGCGCTACTCCATG 181
Db      41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlnIleValSerIleAspSerMet 60
QY      182 CAGGTCTATGAAGGCTTACATCATCATCACCAAGGTTCTGCCAAGAGAGAAATCC 241
Db      61 GlnValIlyrGluGlyLeuAspIleIleThrAsnLysValSerAlaGlnIuGlnArgIle 80
QY      242 TGCCGGCACCATATGATCAGCTTGTGATCCTCTTGACCAATTAACAGTGGTGAC 301
Db      81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnIlyrThrValValAsp 100
QY      302 TTCAGAAATAGCAACTGCTGTGATGAAGATATATTGCCCGAGACAAATTCCTATT 361
Db      101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120
QY      362 GTTGTGGGAGAGAACCAATATTATACATGATCTGCTGTGACCAATTAACAGTGGTAC 421

```

Db 121 ValValGly***ThrasnTyrrTllegluserLeuTrrpLysValLeuValasnThr 140
QY 422 AAGCCCCAGAGATGGGCACTGAGAAAGTATGACCGGAAAGTGGAGCTTGAAGAGAG 481
Db 141 LysProGlnGlu***GlyThrGlnLysValIleAspArgLysValGluLeuGlnLysGlu 160
QY 482 GATGGCTCTGTACTTCACAAAGCGCTTAAGCCAGGTGGAGCCCGAAAGTGGCTCCCAAGCTG 541
Db 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGluMetAlaIleLysLeu 180
QY 542 CATCCACATGACAAAGCGAAGTGGCCAGGAGCTTGCAGAGTTTGGAAAGACAGGAATC 601
Db 181 HisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGluGlnThrGlyIle 200
QY 602 TCTCATAGTGAATTTCTCCATCGTCACATAC-GGAAGAAGGTGGTGTCCCT 654
Db 201 SerHisSerGlnPheLeuHisArgGlnHisThrGlyArgLysTrrpSerPro 218
RESULT 11
US-10-072-326-586
; Sequence 586, Application US/10072326
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P206C1
; CURRENT APPLICATION NUMBER: US/10/072,326
; CURRENT FILING DATE: 2002-02-11.
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 586
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-072-326-586
Alignment Scores:
Pred. No.: 2,06e-95 Length: 222
Score: 1074.00 Matches: 214
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.17% Mismatches: 3
Query Match: 30.04% Indels: 1
DB: 24 Gaps: 0
US-09-513-151-3 (1-2041) x US-10-072-326-586 (1-222)
QY 2 TGGCATAAATGAGCGCGTCCGCGGCTGCACAGCAGTCCCTGGGCACTGGGCTCAGG 61
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValIleSerGlyLeuArg 20
QY 62 GGCCTCAACGAGCGCTTCTAGTATCTCGGGCCGCGGCAACCGGCAATCC 121
Db 21 GlyLeuGlnArgThrLeuProLeuValIleLeuGlnValAlaThrGlyThrGlyLysSer 40
QY 122 ACGCTGGCGCTGAGCTAGGCGGCGGCTCGGAGTGCATCGCTGACGCTGACTCCATG 181
Db 41 ThrLeuAlaLeuGlnLeuGlnArgLeuGlnArgLysGlnIleValSerAlaAspSerMet 60
QY 182 CAGGTATATGAAGCGCTTAACATCATCACCACAAAGTTTCTCCCAAGAGCAGAATC 241
Db 61 GlnValTyrrGlnIleLysAspIleThrAsnLysValSerAlaGlnGlnGlnArgTyrIle 80
QY 242 TGGCGGACACATGATCAGCTTGTGATCCCTGTGACCAATTAACACAGGTGGAC 301
Db 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrThrValValAsp 100

QY 302 TTCAGAAATAGACAACTCCTCTGATGAGATATATTTGCCCGAGCAAAATTCCTAT 361
Db 101 PheAspAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120
QY 362 GTTGGGAGAGAACCAATATTACATTGATATCTGCTCTGGAAGTCTTCTGCAATAC 421
Db 121 ValValGly***ThrasnTyrrTllegluserLeuTrrpLysValLeuValasnThr 140
QY 422 AAGCCCCAGAGATGGGCACTGAGAAAGTATGACCGGAAAGTGGAGCTTGAAGAGAG 481
Db 141 LysProGlnGlu***GlyThrGlnLysValIleAspArgLysValGluLeuGlnLysGlu 160
QY 482 GATGGCTCTGTACTTCACAAAGCGCTTAAGCCAGGTGGAGCCCGAAAGTGGCTCCCAAGCTG 541
Db 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGluMetAlaIleLysLeu 180
QY 542 CATCCACATGACAAAGCGAAGTGGCCAGGAGCTTGCAGAGTTTGGAAAGACAGGAATC 601
Db 181 HisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGluGlnThrGlyIle 200
QY 602 TCTCATAGTGAATTTCTCCATCGTCACATAC-GGAAGAAGGTGGTGTCCCT 654
Db 201 SerHisSerGlnPheLeuHisArgGlnHisThrGlyArgLysTrrpSerPro 218
RESULT 12
US-10-072-326-784
; Sequence 784, Application US/10072326
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P206C1
; CURRENT APPLICATION NUMBER: US/10/072,326
; CURRENT FILING DATE: 2002-02-11.
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 784
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-072-326-784
Alignment Scores:
Pred. No.: 2,06e-95 Length: 222
Score: 1074.00 Matches: 214
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.17% Mismatches: 3
Query Match: 30.04% Indels: 1
DB: 24 Gaps: 0
US-09-513-151-3 (1-2041) x US-10-072-326-784 (1-222)
QY 2 TGGCATAAATGAGCGCGTCCGCGGCTGCACAGCAGTCCCTGGGCACTGGGCTCAGG 61
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValIleSerGlyLeuArg 20
QY 62 GGCCTCAACGAGCGCTTCTAGTATCTCGGGCCGCGGCAACCGGCAATCC 121
Db 21 GlyLeuGlnArgThrLeuProLeuValIleLeuGlnValAlaThrGlyThrGlyLysSer 40
QY 122 ACGCTGGCGCTGAGCTAGGCGGCGGCTCGGAGTGCATCGCTGACGCTGACTCCATG 181
Db 41 ThrLeuAlaLeuGlnLeuGlnArgLeuGlnArgLysGlnIleValSerAlaAspSerMet 60
QY 182 CAGGTATATGAAGCGCTTAACATCATCACCACAAAGTTTCTCCCAAGAGCAGAATC 241

```
Db 61 GlnValTylGluLysLeuAspIleIleThrAsnLysValSerIleGlnGluGlnAlaGlyIle 80
QY 242 TGGCGGACACCATGATGACGTTGTGGATCTCTTGACCAATACACAGTGGTGGAC 301
Db 81 CysAlaHISHisMetIleSerPheValAspProLeuValThrAsnTyrThrValValAsp 100
QY 302 TTCAGAAATAGACACTGCTGATTGAAGATATATTGGCCGACACAAATTTCTATT 361
Db 101 PheAlaAsnAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle 120
QY 362 GTTGTGGAGGAGCAATATTACATTGAATCTGCTGTGAAATCTTGTCAATACC 421
Db 121 ValValIglY**ThrAsnTyrTyrIleGluSerLeuLeuTyrPylValLeuValAsnThr 140
QY 422 AAGCCCCAGAGATGGGACCTGGAAGAGTATTGACCCGAAAGTGGAGCTTGAAGAGAG 481
Db 141 LysProGlnGlu**GlyThrGluLysValIleAspArgLysValGluLeuGluLysGlu 160
QY 482 GATGCTGTGACTTACAAACGCTAGACAGCTGACCCGAAATGGCTGCCAGCTG 541
Db 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaAlaLysLeu 180
QY 542 CATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGATTTTTGAAGAAACAGCAATC 601
Db 181 HisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGluGluThrGlyIle 200
QY 602 TTCATAGAGCAATTTTCCATGCTGTCACATAC-GGAAGAGGTGGTGGCTCCCT 654
Db 201 SerHisSerGluPheLeuHisArgGlnHisThrGlyArgLysTyrPheSerPro 218
```

RESULT 13

```
US-09-757-028-2199
; Sequence 2199, Application US/09757028
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM001
; CURRENT APPLICATION NUMBER: US/09/757,028
; CURRENT FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 2660
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2199
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-757-028-2199
```

Alignment Scores:

```
Pred. No.: 1,47e-57 Length: 143
Score: 686.50 Matches: 131
Percent Similarity: 97.10% Conservative: 3
Best Local Similarity: 94.93% Mismatches: 3
Query Match: 19.20% Indels: 1
DB: 21 Gaps: 1
```

US-09-513-151-3 (1-2041) x US-09-757-028-2199 (1-143)

```
QY 920 AACCAAGCTTTCTAAAGAAAGAGACCTGTCCATTGTCCTGTATGCTTAGAGTA 979
Db 7 AsnAlaGpHeLeuSerArg---ProGlyProIleValProProValTyrGlyLeuGlnVal 25
QY 980 TCTGATGCTCGAAGTGGAGAGTCTGTTTGAACCTGCTTGAATCGTGAAGT 1039
Db 26 SerAspValSerLysTyrIleGlnLysSerValLeuGlnProAlaLeuGlnIleValGlnSer 45
QY 1040 TTCATCCAGGGCCACAGGCTACAGCCACTCCAAATGAAGATGCATACATGAAGTGA 1099
Db 46 PheIleGlnGlyHisLysProThrAlaThrProIleLysMetProTyrAsnGlnLysGlu 65
```

```
QY 1100 AACCAAGAGATTATACCTGTGTGACCTGTGTGATGCAATCATCATGGGATGGCGAA 1159
Db 66 AsnLysArgSerTyrHisLysLeuCysAspLeuCysAspArgIleIleIleGlyAspArgGlu 85
QY 1160 TGGCGGACGACATTAATATCCAAATCCACTTGAACCACTGAAGAAAGAGAGATTG 1219
Db 86 TrpAlaAlaHisIleLysSerLysSerHisLysAsnGlnLeuLysLysArgArgLeu 105
QY 1220 GACTCAGATGCTGTCAACACATAGAAAGTCAAGTGTTCCTCCAGACTATACAAAGAA 1279
Db 106 AspSerAspAlaValAsnThrIleGluSerGlnSerValSerProAspHisAsnLysGlu 125
QY 1280 CCTAAAGGAGAGGATCCCGAGGCGCAATGATCAAGACTGAATGACGCTT 1333
Db 126 ProLysGlnLysGlySerProGlyGlnAsnAspGlnGluLeuLysCysSerVal 143
```

RESULT 14

```
US-10-222-911-2199
; Sequence 2199, Application US/10222911
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM001CIN
; CURRENT APPLICATION NUMBER: US/10/222,911
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 09/757,028
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 2660
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2199
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-911-2199
```

```
Alignment Scores:
Pred. No.: 1,47e-57 Length: 143
Score: 686.50 Matches: 131
Percent Similarity: 97.10% Conservative: 3
Best Local Similarity: 94.93% Mismatches: 3
Query Match: 19.20% Indels: 1
DB: 26 Gaps: 1
```

US-09-513-151-3 (1-2041) x US-10-222-911-2199 (1-143)

```
QY 920 AACCAAGCTTTCTAAAGAAAGAGACCTGTCCATTGTCCTGTATGCTTAGAGTA 979
Db 7 AsnAlaGpHeLeuSerArg---ProGlyProIleValProProValTyrGlyLeuGlnVal 25
QY 980 TCTGATGCTCGAAGTGGAGAGTCTGTTTGAACCTGCTTGAATCGTGAAGT 1039
Db 26 SerAspValSerLysTyrIleGlnLysSerValLeuGlnProAlaLeuGlnIleValGlnSer 45
QY 1040 TTCATCCAGGGCCACAGGCTACAGCCACTCCAAATGAAGATGCATACATGAAGTGA 1099
Db 46 PheIleGlnGlyHisLysProThrAlaThrProIleLysMetProTyrAsnGlnLysGlu 65
QY 1100 AACCAAGAGATTATACCTGTGTGACCTGTGTGATGCAATCATCATGGGATGGCGAA 1159
Db 66 AsnLysArgSerTyrHisLysLeuCysAspLeuCysAspArgIleIleIleGlyAspArgGlu 85
QY 1160 TGGCGGACGACATTAATATCCAAATCCACTTGAACCACTGAAGTGAAGAAAGAGATTG 1219
Db 86 TrpAlaAlaHisIleLysSerLysSerHisLysAsnGlnLeuLysLysArgArgLeu 105
QY 1220 GACTCAGATGCTGTCAACACATAGAAAGTCAAGTGTTCCTCCAGACTATACAAAGAA 1279
Db 106 AspSerAspAlaValAsnThrIleGluSerGlnSerValSerProAspHisAsnLysGlu 125
```


OY 1280 CCTAAGGGAAGGATCCCGAGGAGATGATCAAGCTGAAATGACGCGT 1333
Db 126 ProlysglyglyserProglylnasnaspglnglnleuLycyserval 143

RESULT 15

US-09-270-767-44570
; Sequence 44570, Application US/09270767
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44570
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-44570

Alignment Scores:

Pred. No.:	1,01e-48	Length:	252
Score:	597.50	Matches:	119
Percent Similarity:	67.48%	Conservative:	47
Best Local Similarity:	48.37%	Mismatches:	77
Query Match:	16.71%	Indels:	3
DB:	16	Gaps:	2

US-09-513-151-3 (1-2041) x US-09-270-767-44570 (1-252)

OY 71 CGGACCTACCTCTGTGATGATCTCGGGGCGAGCGGCGCAATTCACGCTGCG 130
Db 5 ArglyValProleuileValValLeuGlyserThrGlyThrGlyThrLysLeuSer 24
OY 131 TTGCGAGCTAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 190
Db 25 LeuGlnLeuAlaGlnArgPheGlyGlyGlyLeuIleSerPheLysPheMetGlnValTyr 44
OY 191 GAAGGCTAGACATCATCACAACAGCTTTCGCCAGAGCAGAGCAATTCGCCGCGAC 250
Db 45 ThrHisLeuAspIleAlaThrAlaLysAlaThrLysGlnGlnSerArgAlaArgHis 64
OY 251 CACATGATCAGCTTTGTGGATCCTTGTGACCAATTACACAGTGGTGCATTGAGAAT 310
Db 65 HisLeuLeuAspValThrThrPro--AlaGlnProPheThrValThrHisPheArgAsn 83
OY 311 AGAGCACTGCTGTGATGATATATTTGGCCGAGCAAAATTCCTATTTGTGGCA 370
Db 84 AlaAlaLeuProIleValGlnArgLeuAlaLysAspThrSerProIleValValGly 103
OY 371 GGAACCAATTTATCATTTGATGATCTGCTGGAAGTCTTGTCAATAC-----AAG 424
Db 104 GlyThrAsnTyrTyrIleGlnSerLeuLeuTrpAspIleLeuValAspSerAspValLys 123
OY 425 CCCGAGAGATGGGCACTGAGAAAGTATTGACCGGAAAGTGCAGCTGAAAAAGAGAGAT 484
Db 124 ProGlnGlnGlyLysProSerGlyLysHisLeuLysAspAlaGlnLeuAsnAlaLeuSer 143
OY 485 GGTCTTTGACTTCACAAAGCGCTAAGCCAGGTGGAGCCGAAATGCGTCCAAGCTGCAT 544
Db 144 ThrLeuGlnLeuHisGlnLeuLeuAlaLysIleAspAlaGlySerAlaAsnArgIleHis 163
OY 545 CCACATGACAAAGCAAGGCGGCGAGCTTGCAAGTTTGAAGAACAAGCAAGATCTCT 604
Db 164 ProAsnAsnArgArgLysIleIleArgAlaIleGlnValTyrGlnSerThrGlyGlnThr 183
OY 605 CATAGTGAATTTCTCCATGCTCAACATAGCAAGAGTGTGTGCTCCCTTGAGGTCTCT 664
Db 184 LeuSerGlnMetLeuAlaGlnArgAlaGlnProGlyGlyAsnArgLeuGlyGlyPro 203
OY 665 CTGAAGTTCTTAACCTTGACATCCTTGGCTTCATGCTGACCAGGCACTTCAATGAG 724

Db 204 LeuArgTyrProHisIleValLeuLeuTrpLeuArgCysGlnGlnAspValLeuAsnGln 223
OY 725 CCGTTGATTAAGAGGCGGATGATGACATGCTGCTGCGGCGCTTGTGAGAACTAAGAGAT 784
Db 224 ArgLeuAspSerArgValAspGlyMetLeuAlaGlnGlnLeuLeuProGlnLeuArgGln 243
OY 785 TTTCACAGACGCTATAAT 802
Db 244 PheHisAsnAlaHisHis 249

Search completed: April 21, 2003, 19:12:17
Job time : 321.711 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW nucleic - protein search, using frame_plus.n2p model

Run on: April 21, 2003, 18:41:35 ; Search time 22.0286 seconds
(without alignments)
5452.195 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 3575
Sequence: 1 CCGCCATAGATGCGCTCCG.....TTTACAGAAAAA 2041

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Zgapop 6.0, Zgapext 7.0
Delop 6.0, Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xih
-Q/cgn2_1/USPTO_spool/US09513151/runat_15042003_141144_26406/app_query.fasta.1.2446
-DB=Issued_Patents_AA -OFMT=fasta -SUFFIX=n2p.rai -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US09513151 @cgn_1.1.23 @unat_15042003_141144_26406 -NCPU=6 -ICPU=3
-NO.XLTRY -NO.MMAP -LARGEOUTRY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/6C.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/6D.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307.5	8.6	193	4	US-09-134-001C-4171
2	114.5	3.2	563	4	US-09-134-001C-3172
3	99	2.8	1024	4	US-09-562-737-50
4	99	2.8	1548	1	US-08-463-092B-7
5	99	2.8	1548	2	US-08-460-907B-7
6	98.5	2.8	1044	2	US-08-777-405A-2
7	98.5	2.8	1044	2	US-08-977-871A-2
8	98.5	2.8	1044	2	US-09-225-951-2
9	97	2.7	511	2	US-09-019-201A-2
10	96.5	2.7	430	4	US-09-310-363C-6
11	95.5	2.7	3052	2	US-08-557-122A-26
12	95.5	2.7	3052	4	US-09-262-666-26

13	94	2.6	529	2	US-08-836-442-3	Sequence 3, Appl1
14	94	2.6	566	4	US-09-040-725A-1	Sequence 1, Appl1
15	94	2.6	739	2	US-08-836-943-2	Sequence 2, Appl1
16	94	2.6	859	4	US-09-199-637A-281	Sequence 281, App
17	93.5	2.6	236	4	US-09-605-858-34	Sequence 5, Appl1
18	93.5	2.6	240	4	US-09-073-887-5	Sequence 5, Appl1
19	93	2.6	674	4	US-08-961-083-200	Sequence 200, App
20	93	2.6	633	4	US-08-235-836C-72	Sequence 72, Appl1
21	92.5	2.6	544	3	US-08-559-397A-29	Sequence 29, Appl1
22	92.5	2.6	3075	2	US-08-460-309-5	Sequence 5, Appl1
23	92.5	2.6	3075	2	US-08-125-077-5	Sequence 5, Appl1
24	92.5	2.6	3696	4	US-09-134-001C-5080	Sequence 5080, Ap
25	91.5	2.6	976	4	US-09-104-324B-4	Sequence 4, Appl1
26	91.5	2.6	2482	1	US-08-328-254-6	Sequence 6, Appl1
27	91	2.5	1150	4	US-09-002-285-74	Sequence 74, Appl1
28	90.5	2.5	429	2	US-09-074-512-1	Sequence 1, Appl1
29	90.5	2.5	510	4	US-09-134-001C-3368	Sequence 3368, Ap
30	90.5	2.5	649	4	US-09-134-001C-3891	Sequence 3891, Ap
31	90.5	2.5	693	4	US-08-235-836C-68	Sequence 68, Appl1
32	90.5	2.5	1201	4	US-09-098-901-2	Sequence 2, Appl1
33	90.5	2.5	1872	1	US-08-188-582-14	Sequence 14, Appl1
34	90.5	2.5	1872	1	US-08-646-715-14	Sequence 14, Appl1
35	90.5	2.5	1893	1	US-08-188-582-11	Sequence 11, Appl1
36	90.5	2.5	1893	1	US-08-646-715-11	Sequence 11, Appl1
37	90	2.5	425	4	US-09-310-363C-8	Sequence 8, Appl1
38	90	2.5	474	4	US-09-134-001C-4953	Sequence 4953, Ap
39	90	2.5	630	4	US-09-360-545-78	Sequence 78, Appl1
40	90	2.5	655	1	US-08-264-002-2	Sequence 2, Appl1
41	90	2.5	1384	4	US-08-976-255-11	Sequence 11, Appl1
42	89.5	2.5	1262	4	US-09-357-251-33	Sequence 33, Appl1
43	89.5	2.5	1266	1	US-08-468-557-4	Sequence 4, Appl1
44	89.5	2.5	1266	4	US-09-357-251-32	Sequence 32, Appl1
45	89.5	2.4	2511	4	US-09-261-907-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-134-001C-4171
Sequence 4171, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NOCLEDIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064, 964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055, 779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4171
LENGTH: 193
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4171

Alignment Scores:

Pred. No.: 1.52e-25
Score: 307.50
Percent Similarity: 61.60%
Best Local Similarity: 36.53%
Query Match: 8.60%
DB: 4
Gaps: 3

US-09-513-151-3 (1-2041) x US-09-134-001C-4171 (1-193)

QY 68 CAAGGACCCCTACTCTCTTGTAGTGAATTCGGGGCCACGCGCAATTCACAG 124
DB 3 GlnMetThrLysPhePheValIleValIleGlyProThrAlaSerIlyLysThrIle 22


```

Db 365 -----Ileasplileglyllysleusenglnluarqargayalaiaarggluleu 382
OY 1163 GCAGCCCATATAAATCCAAATCCACTTGACCAACTGAAGAAAAG----- 1210
Db 383 ArgasphsilelvalsergluileglnasleuGlnmetlysasprlaasleuGluile 402
OY 1211 -----AGAAGATTGACCTCAGATGCTGTCACACACATAGAAAGTCAGAGT-----GTT 1258
Db 403 Serphelysproleusargpuprothrllegllyllegluphevalglupheleuile 422
OY 1259 TCCCCAGACTATACAAAGAACCTTAAGGAGGATGCCCA 1300
Db 423 Serproasn-----ArgGlyGluPro 429

RESULT 3
US-09-562-737-50
; Sequence 50, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSM0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-50

Alignment Scores:
Pred. No.: 0.0913 Length: 1024
Score: 99.00 Matches: 84
Percent Similarity: 40.16% Conservative: 63
Best Local Similarity: 22.95% Mismatches: 159
Query Match: 2.77% Indels: 60
DB: 4 Gaps: 19

US-09-513-151-3 (1-2041) x US-09-562-737-50 (1-1024)
OY 293 GTGTGAGACTTCAGAAATAGACCACTGCTGATGAAGATATATTGCCGAGACAAA 352
Db 101 MetValAspAspserlystrValThrAsnMetLeuMetThrIleCysAlaArgIleGln 120
OY 353 ATTCTATTGTTGTGGAGAGAACCAATTATTAATTGAACTCTGCTGTGGAAGTTCTT 412
Db 121 Ile-----ThrAsnHisAspIleuTrpSerLeuArgGluLeuMet 134
OY 413 GTAATAATCAAGCCCAAGAGATGGGCACT-----GAGAAAGTGAT----- 454
Db 135 GluGluLysLysgluserIleThrGlyThrLeuArgLysAspLysValLeuLeuArgAsp 154
OY 455 GACCGAAAGTAGGAGCTTGAAGAGGAGGATGCTTGTACTGTCC-----AAGCGCTA 508
Db 155 GluLysLysMetGluTrpLeuLysGlnLys-----LeuHisTrpAspAspTrpLeu 171
OY 509 AGCCAGGTGAGCCAGAAATAGCTGCCAAG-----CTGCATCCACATGACAAACGC 559
Db 172 AsnTrpLeuAspHisGlyArgThrAlaArgGluGlnLysAlaGluGlnLysGluAspLeu 191
OY 560 AAGATGGCCAGAGACTTGCAGATTTTGAAGAAAACAGAAATCTCTCATAGTAATT--- 616
Db 192 LeuLeuArgArgLysPhePheTyrgLysGlnAsnValAspSerArgAspProPheGln 211
OY 617 CTCACATCGTCAACATACGGAAGAGGTGGTGGTCCCTTGAGAGT-----CCTCTGAG 670
Db 212 LeuAsnLeuLeuTyrgAlaGlnAlaGlyAspAspIleLeuAsnGlySerHisProHisSer 231

```

```

OY 671 TTCTTAACCCCTTGACCTCTTGGCTTCATAGCTGACACGAGCACTTCTAGATAGCCCTTG 730
Db 232 PheAspLysAlaLysglupheleuAlaIlePheGlnCysGlnIleGlnPheGlyProLysAsn 251
OY 731 GATAGAGAGGTGATGACATGCTTGGCTGCGGGCTCTGGAGAGACATGAGATTTTTCAC 790
Db 252 GluGlnLys-----HisLysAlaGlyPheMet---AspLeuLysAspPheLeu 266
OY 791 AGACGCTATATACAGAAATGTTTCGGAATAATACCGAGCATATCAAT----- 841
Db 267 ProLysGlnAsnValLysGlnLysGluArgLysIleGlnIleHisLysAsnCys 286
OY 842 GGTATCTTCCAAATCAATGCTTTCAGAGAAATTTACAGAGTACCTGATCATGAGGAAA 901
Db 287 GlyGlnMetArgGluIleGlnAlaLysValArgTyrgValSerLeuAlaArgSerLeuLys 306
OY 902 TGCACACTGAGACTAGTACCAAGCTCTTAAGAA-----GGACCTGGTCCCAT 952
Db 307 ---ThrTyrgLysThrSerPhePheLeuValLysGlnLysMetValGlyLysAsnLysLeu 325
OY 953 GTCCCCCTGCTATAGCTTATAGAGTATCTGATGCTCTGAAATGCGAGAGAGCTGTTCTT 1012
Db 326 ValProArgLeuTrpGlyIleThrLysGluCysValMetArgTyrgLysGluLys----- 343
OY 1013 GAACCTGCTTGAATCGCAAGATTTC-----ATCCAGGCCACAAAGCTTACAGCC 1066
Db 344 -----ThrLysGluValIleGlnAlaTrpAsnLeuThrAsnIleLysArgTrpAspAla 361
OY 1067 ACTCCATTAAGATGCCATCAATGAGCTGAGACAGAAAGAGTTTACCTGCTGTGAC 1126
Db 362 SerProLysSerPheThrLeuAspGluGlyAspTyrgLysGlnAspGlyTyrgSerPheGln 381
OY 1127 CTCGTGAT-----CGAATCATCATTTGGGATCGGAGTGGCAGAGCCAGCC 1171
Db 382 ThrThrGluGluGlnGlnIleAlaGlyLeuIleAlaGlyLysLeuAsp---IleIleHis 400
OY 1172 ATAAATCCAAATCCCACTTGACCAACTGAAGAAAGAAAGATTTGAGTCAAGATGCT 1231
Db 401 LysLysLysLysSer-----LysAspHisPheIleLeuGlnGlyAspGlu 415
OY 1232 GTCACACCATAGAAAGTCAAGAGTGTTCCTCCCA-----GAC 1267
Db 416 GluSerThrMetLysGluAspSerValSerProLysLysSerLeuValLeuGlnGlnGln 435
OY 1268 TATAACAAGAACCTTAA 1285
Db 436 TyrAsnArgValMetLys 441

RESULT 4
US-08-463-092B-7
; Sequence 7, Application US/08463092B
; Patent No. 5766880
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTEO RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; Application Number: US/08/463,092B

```



```

1 PRIOR APPLICATION DATA:
2 APPLICATION NUMBER: 08/029,340
3 FILING DATE: 8-MAR-1993
4 CLASSIFICATION: 424
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER: 08/141,893
7 FILING DATE: 26-OCT-1993
8 CLASSIFICATION: 424
9 PRIOR APPLICATION DATA:
10 APPLICATION NUMBER: 08/407,207
11 FILING DATE: 20-MAR-1995
12 CLASSIFICATION: 424
13 ATTORNEY/AGENT INFORMATION:
14 NAME: Steeg, Carol Mternickl
15 REGISTRATION NUMBER: 39,539
16 REFERENCE/DOCKET NUMBER: 01551
17 TELECOMMUNICATION INFORMATION:
18 TELEPHONE: (613) 545-2342
19 TELEFAX: (613) 545-6853
20 INFORMATION FOR SEQ ID NO: 7:
21 SEQUENCE CHARACTERISTICS:
22 LENGTH: 1548 amino acids
23 TYPE: amino acid
24 TOPOLOGY: linear
25 MOLECULE TYPE: peptide
26 FRAGMENT TYPE: Internal
27
28 OS-08-460-907B-7

```

Alignment Scores:	
Pred. No.:	0.12
Score:	99.00
Percent Similarity:	36.65%
Best Local Similarity:	21.51%
Query Match:	2.77%
DB:	2
Length:	1548
Matches:	74
Conservative:	52
Mismatches:	88
Indels:	130
Gaps:	19

US-09-513-151-3 (1-2041) x US-08-460-907B-7 (1-1548)

OY	14	GGCGCCGTGGCGCTGCACAGCAGATTCCTGTGGGACAGTGGG-----CTCAGGGGC	64
Db	1271	AlaGlyAlaAlaSerSerAlaLeuHisProValGlnAlaGlySerLeuValLeuGluGly	1250
OY	65	CTGCAG-----CGGACCTACCTTTGAGTG-----	91
Db	1391	ValGlnMetArgTyrArgGluGlyLeuValProLeuValLeuArgValSerPheGlnIle	1310
OY	92	-----ATTCTGGGGGCGGCGGACGGGAAATCCAGCTGGCG	130
Db	1311	AlaProArgGluIuysValGlyIleValGlyArgThylGlySerGlyIysSerThrIleuLeu	1330
OY	131	TTTCAGCTAGCGCAGCGGCTGGCGGGTGAATCGTCAGCGCTCAGCTCAGAGGTCTAT	190
Db	1331	LeuThrPhe-----MetArgMetVal	1337
OY	191	GAA-----GGCCTAGACATCATCCACCAC-----AAGGTTTTCGCCAAGACAGCAGGA	238
Db	1338	GluValGlySerGlyValIleHisValAsnGlyArgGluMetSerAlaTyrGlyLeuArg	1357
OY	239	ATTCGGCGGACACATGATCATCTTTGTG-----GATCCCTTGTGACCAATTATACAA	292
Db	1358	GluLeuArgArgHis---PheSerMetIleProGlnAspProValLeuPheAspIlyThr	1376
OY	293	GTC-----	295
Db	1377	ValArgGlnAsnValAspProPheLeuGlnAlaSerSerAlaGluValIleTrrAlaLeu	1396
OY	296	-----GTGACCTTCAGAAATATAGACCAACTGCTCTGATTGAGATATATTTGCCGAGAC	349
Db	1397	GluLeuValGlyLeuArgGluArgValAlaSerGluSerGluGlyIleLeuSerArg---	1415
OY	350	AAAATTCCTATGTGTGGGAGGAGACCAATTAATTACATTGAATCTCTGCTGGAAAGTT	409
Db	1416	-----ValLeuGluGlyGlySerAsnTyrSerVal-----	1429

QY	410	CTTGTCAATCAATCAAGCCGCCAGAGGAGGGGACAGTGAAGAAATGATTCACGAAAAAGTGAG	466
Db	1426	-----gLyInIaRgInIleuMetCySMetIaIaRgIaLeu-----	1437
QY	470	CTTGAAGAAAGAGAGATGGCTTTGATTT---CACAAAGCCTAAGCCAGGTGGACCCAGAA	526
Db	1438	LeuIysArGgIySerGIyRheIleuMetArGgIaIaThIaAsnIleAsProIaI	1457
QY	527	ATGGCTGCCAAGTGCATGCATGACGAAACGAAAGTGGCCAGAGGCTTGGAAATTTT	586
Db	1458	Leu-----AaRgIaGInIleGInaIaThIaValMetSerIaIaRhe	1470
QY	587	GAAAGAACAGAAATCTCTCATAGTGAATTTCTGCATGCTGAACATAGAGAAAGGTGT	646
Db	1471	-----SerIaTgThValIleThIaIleThIaIaIaIaRgLeuInIstIhValI-----	1485
QY	647	GGTCCCTTGGAGGTGCTGTGAAGTTCTGAACCTTGCATGCTTTGGCTTCATGTGAC	706
Db	1485	-----	1485
QY	707	CAGGACATCTAAGTAGAGCCCTGGATGAAGAGGTGATGACATGCTTGGCTGGGCTG	766
Db	1486	-----AlaGInIyRAsRySIleIle-----ValMetArpIaSIyIaI	1498
QY	767	TTTGAGAGAACTAAGAGATTTTCAACAAACGCTTAATCAAGAAAGTTCGAAAAATGAC	826
Db	1499	ValIaIaIuMetGly-----SerProIaRgIuIeuValMetAsIhIst	1512
QY	827	CAGAGCATATCAACATGATATATTCATCAATCAATT-----GGCTTCAAGAAATTCAC	877
Db	1513	GInSerMetRheIstSerMetValGIuSerLeuGIySerArGIySerLyAsRheIyR	1532
QY	878	GAGTACTGATC 889	
Db	1533	GIuIeuIeuMet 1536	

RESULT 6

US-08-777-405A-2
; Sequence 2, Application US/08777405A

GENERAL INFORMATION:

APPLICANT: Chantry, David
APPLICANT: Hoekstra, Merl

APPLICANT: Holtzman, Douglas A
TITLE OF INVENTION: No. 5858753e1 Lipid Kinase

```

;
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: MARS
STREET: 6300 SE

; CITY: Chicago
; STATE: Illinois
;; COUNTRY: USA
; ZIP: 60606
;

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
;

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, v
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/777,405A
FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: No. 5858753and, Greta E.
REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 344-1111
TELEFAX: (312) 344-1111

TELEX: 25-3856
INFORMATION FOR SEO ID NO: 2

SEQUENCE CHARACTERISTICS

REFERENCE/DOCKET NUMBER: 27866/33441
TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1044 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-225-951-2

Alignment Scores:

Pred. No.: 0.105

Score: 98.50

Percent Similarity: 34.258

Best Local Similarity: 20.308

Query Match: 2.764

DB: 2

Length: 1044

Matches: 96

Conservative: 66

Mismatches: 174

Indels: 137

Gaps: 21

US-09-513-151-3 (1-2041) x US-09-225-951-2 (1-1044)

```

OY 291 CAGTGTGACTTCAGAAATAGACCACTGCTGATTGAGATATATTGCCGAGACA 350
DB 10 GlnpHePrThrLysGlnGlnuAsnGlnSerValValValAspPheLeuLeuProThrgly 29
OY 351 AAATTCCTA-----TTCTTG 365
DB 30 ValTYrLeuAsnPhProValSerArGAsnAlaAsnLeuSerThrLLeysGlnLeuLeu 49
OY 366 TGGGAG-----GACCAATTATTACATGATCTGCTGCTGAGAAAT-----408
DB 50 TrpHisArGAlaGlnTyGlnPro-LeuPheHisMetLeuSerGlyProGlnAlaTyVa 69
OY 409 -----TC 410
DB 69 LpHeThrcysIleAsnGlnThrAlaGlnGlnGlnuLeuGlnuAspGlnuArgyle 89
OY 411 TTGCTAATCCAA-----GCCCGAGAGATGGCA 440
DB 89 uCysAspValGlnProPheLeuProValLeuArgLeuValAlaArgGlnLysP-Argy 109
OY 441 CTGAGAAAGTGTATGACCAAGAGTGGCTTGTGACTGAC 500
DB 109 aLlysLysLeuLeuAsnSerGlnLLeSerLeuLeuLLeGlyLysGly-----LeuHisG 127
OY 501 AACGCTAAGCCAGGTGAGCCAGAAATG-----GCTCCAGCTGCATCCACATG 551
DB 127 LpHeAspSerLeuCysAspProGlnValAlaAsnAspPheArgAlaLysMetCysGlnPhec 147
OY 552 ACAAAAGCAAGTGGCCAGGAGCTTGCAGTTTGAAGAAACAGAGATCTCATAGTG 611
DB 147 ySglnGlnuAlaAlaArgArgGlnGlnuLLeGlyTrpGlnuAlaTrpLeuGlnuTySerP 167
OY 612 AATTTCCTCATGCTCAACTAGCGAAGAGTGCTGCTGCTGAGAGTCTGTAAT 671
DB 167 heProLeuGlnuLeuLpProSerAlaGlnuThrTrpLysPro-----GlyThrLeuArgL 185
OY 672 TCTTAACCTTGCATCTT-----TGCTTCATGCTGCAGGAGAGTCTAG 719
DB 185 euProAsnArgAlaLeuLeuValAsnValLysPheGlnuLysSerGlnuLysPheThP 205
OY 720 ATGAGCGCTTGATAGAGGTGATGATGATCTTCTGCT-----760
DB 205 heGlnValSerThrLysAspValProLeuAlaLeuMetAlaCysAlaLeuArgLysLysA 225
OY 761 -----GGCTCTTGAGAGAACTAAGATTTTCAACACCTAATC 803
DB 225 laThValPheArgGlnProLeuValGlnuLpProGlnuAspTyThrLeuGlnuValAsnG 245
OY 804 AGAAGATGTTTGGAAATATAGCCAGGACTATCAACATGTTTCCATCATCAATGGCT 863

```

```

DB 245 LysrHis-----GluTYrLeuTYrGlyAsnTYrProLeuCysGlnP 259
OY 864 TCAAGAA-----TTTACAGACTACCTGATCACTGAGGAAATATGCACCTGG 911
DB 259 heGlnTYrLLeCysSerCysLeuHisSerGlyLeuThrProHis-----LeuThMeV 277
OY 912 ACAGTAGTACCACTTCA-----AAGAAAGACCTGGTCCCATGTC- 955
DB 277 aLHisSerSerLLeuAlaMetArGAspGlnuGlnuSerAsnProAlaProGlnValG 297
OY 956 -----CCCCCTGTATGCTTGAAGTATGATGATGCTGCAAGT 995
DB 297 LpLysProArgAlaLysProProPheProAlaLysLysProSerSerValSerLeuT 317
OY 996 GGGAGAGTCTGTTCTTGAACCTGCTTGAATCTGCAAACTTCAACAGGCGCA 1055
DB 317 rPser-----LeuGlnuLpProPheArgyle-----GluLeuLLeGlnuLysL 332
OY 1056 AGCCTACAGCCACTCCA-----ATAAGATGCCATACAGTGAAGCTGAGAACAGAAAGT 1112
DB 332 ySValAsnAlaAspGlnuArgMetLysLeuValValGlnuAlaGlyLeuPheHisGlyAsnG 352
OY 1113 ATCACTGT-----GACCTGTGTATGCAATCATCATTTG 1148
DB 352 LmetLeuCysLysThrValSerSerSerGlnuLysSerGlnuProVal-----370
OY 1149 GGGATCGCAATGGGAGCGGCACATTAATCCAAATCCACTTGAACCACTGAAGAA 1208
DB 371 -----TrpLysGlnuArgLeuGlnuPheAspLLeAsnLLeCysAspLeuProArg 387
OY 1209 GAAGAGATTGAGCTCAGATGCTGTCACACCATAGAAAGTGAAGTGTGTTCCAGACT 1268
DB 387 eAlaLArgLeuCysPheAlaLeuTYrAlaValLLeGlnuLysAlaLysLysAlaArgSerT 407
OY 1269 ATAACAAGAACCTTAAGGAGGAGATCCCGAGGAGATGATCAAGACTGAATGCA 1328
DB 407 hTrLysLysSerLysLys-----413
OY 1329 GCCTTTAAGACATATCCAGTGGCTTGGAAAGTGGTGGGATCCAGTTCCAG- 1384
DB 414 -----AlaAspCysProLLeAlaTrpAlaAsnLeuMetLeuPheAspTYrLysAspG 431
OY 1385 -----AGGAGGGGATGTTGTCTCCAGTGG 1414
DB 431 LpLeuLysThrGlnuArgCysLeuTYrMetTr 442

```

RESULT 9
US-09-019-201A-2
Sequence 2, Application US/09019201A
Patent No. 5968780

GENERAL INFORMATION:
APPLICANT: FENG, PING
APPLICANT: SOPPET, DANIEL R.
APPLICANT: LI, YI
APPLICANT: DILLON, PATRICK J.
TITLE OF INVENTION: DENDRITIC CELL-DERIVED GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/019, 201A
FILING DATE: HEREWITH
CLASSIFICATION: 536


```

ATTORNEY/AGENT INFORMATION:
NAME: A. ANDERS BROOKES
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF346
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 511 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-019-201A-2

Alignment Scores:
Pred. No.: 0.0961 Length: 511
Score: 97.00 Matches: 75
Percent Similarity: 37.73% Conservative: 48
Best Local Similarity: 23.01% Mismatches: 106
Query Match: 2.71% Indels: 97
DB: Gaps: 16

US-09-513-151-3 (1-2041) x US-09-019-201A-2 (1-511)

QY 371 GGAAACCAATTATTCATGTAATCTGCTGGAAAGTTCTTCAATACCAAGCCAG 430
DB 25 GYSERIALeuserileseriluglthrargalanileuleuleuylsgllysmet 44
QY 431 GAGATGGGACGTCAGAAAGATGATGACCCAAAGTGGCTTGAAGAGGATGGCTT 450
DB 45 ARGLEUGLYLARGLEUVALLEUASNTHLYSGLEUGLEUVALASNGLIATGLEUMET 64
QY 491 GTACTTCACAAAGCCGTAAGCCAGTGGACCCAGAAATGGCTGCCAATGCCA--- 547
DB 65 THIRLEU-----LYSILELAGIUMETLYSGLIAMETKRGTHIRLEUIERPROPRO 82
QY 548 -----CATGACAAAGCCAAAGTGGCCAGAGCTTGCAAGTTT--- 586
DB 83 SERMETHISRPHERHEGIALALYSNLSLEULIEGLIADSER---GLNVALIPHEASNILE 101
QY 587 -----GAAACAAAGGAACTCTGCAATAGT 610
DB 102 LEUATGMEETRPOLYSGLIALLALEUHNISLEUHNISAPRIIEGLIYILEVALTHIMET 121
QY 611 GAATTTCTC-----CATGCTCAACATACGGAAGAGTGGTGGCCCTT 655
DB 122 ASPTRPLEUVALARGASNVALTHRTYARGPROHISCYSHISILECYSRHEIHPROARG 141
QY 656 GGA---GGTCCCTGTAAGTTCTTAACCT-----TGCATCCTT 691
DB 142 GLYLIMETGLINPHEARGPHEALANISPROTHRPORARGPROSERGIULYSCYSELYS 161
QY 692 TGGCTTCATGCTGACACAGGAGCTTGAAGAGCGCTTGATGAAGAGGGG--- 742
DB 162 TRP-----LILEULEUGLIASPTYTRGLYSALGVALGILASNVAL 175
QY 743 -----GATGACATGCTGCTGCTGGGCTCTTGAGAGAACTAAGAGATTTTACAGAGA 793
DB 176 THIRLIPHEASPRSERLEU-----LEUATGASNPHEIHPHEU 188
QY 794 CGCTATATTCAGAAAGATGTTTGGAAATATGACGAGACTATGCAACATGGTATCTTCAA 853
DB 189 VALTHIRLHNISPROGLIUALILETYTRHASNGLINASNVALVALTRPSELYSPHEGLI 208
QY 854 TCATTTGGCTC----- 865
DB 209 THIRLIPHEAHPHRTLESERGLYLEUHNISLYTALAPROVALPHEARGASPTYTRVAL 228
QY 866 -----AAGGAATTTACAGATACCTGATCAGTACGAGGAATACACACATGAG 913
DB 229 PHEARGSERMETGLINLIPHEIYTRGLI-----ASPASNVALLEUITYRMEGLI 244

```

```

QY 914 ACTAGTACCAAGCTCTCTAAGAAAGACCTGGTCCCATTTGCCCTGTATGGCTTA 973
DB 245 ILEARGALARGLEU-----PROVALTYR----- 253
QY 974 GAGGTATCTGATCTCTCGAAGTGGAGAGCTGTCTTGTGACCTGCTGTGAATCGTG 1033
DB 254 GLINLEUSERGLYLIHNSHISASRGLUGITRPSERVALYSHIRYTRGINSIUALALA 273
QY 1034 CAAGTTTCATCCAGGCGCCAAAGCCTACAGCCATCAATTAAGATGCCATCAATGA 1093
DB 274 GLNLYSRPHEVALGLIURHNS---PROGLIPHEILEGLYILEYSHILETYRSEAR 292
QY 1094 GCTGAGACACAGAAAGTTTCACTGTGTGACCTCTGTATGATCATCATTTGGGAT 1153
DB 293 HIS-ARGSERLYSASPYAL-----ALAVALLIIEAGIUSERILEARGME 307
QY 1154 CGCGAATGGCGACGCG 1169
DB 307 TALAMETGLYLEUARG 312

RESULT 10
US-09-310-363C-6
Sequence 6, Application US/09310363C
Patent No. 6388169
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
APPLICANT: Shi, Jinrui
APPLICANT: McELIVER, John
APPLICANT: Bowen, Benjamin
APPLICANT: Baszczyński, Christopher
TITLE OF INVENTION: RecA cDNAs and Uses Thereof
FILE REFERENCE: 0855
CURRENT APPLICATION NUMBER: US/09/310,363C
CURRENT FILING DATE: 1999-05-12
PRIOR APPLICATION NUMBER: 60/099,765
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/096,492
PRIOR FILING DATE: 1998-08-14
PRIOR APPLICATION NUMBER: 60/088,529
PRIOR FILING DATE: 1998-06-08
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 430
TYPE: PRT
ORGANISM: Zea mays
US-09-310-363C-6

Alignment Scores:
Pred. No.: 0.0974 Length: 430
Score: 96.50 Matches: 77
Percent Similarity: 33.85% Conservative: 53
Best Local Similarity: 20.05% Mismatches: 135
Query Match: 2.70% Indels: 119
DB: Gaps: 17

US-09-513-151-3 (1-2041) x US-09-310-363C-6 (1-430)

QY 32 CGAGCAGTCTCTGGGCGAGT-----GGGCTCAGG 61
DB 97 ARGASPYALPROVALVALSERTHYSGRPHLEALALEUASPMETALALEUGLYTHIRGLY 116
QY 62 GGCTTGCACAGCCCTTACTCTGTAGTATCTCGGGCCACAGGCGCCCAATCC 121
DB 117 GLYLEUPROLYSGLY---ARGVALILEGLIUALTYGLYPROGLIUALASERGLYSTR 135
QY 122 ACGCTGGCTTGACGCTA-----GGCCAGCGCTCGCGGTGAGATGTCAGCGCT 172
DB 136 THIRLEUALLEUHNISVALIIEAGIUALIAGIULYSASNGLYLYTRYSALAPHEVAL 155
QY 173 GACTCCATGACG-----GCTATGAAGGCTTACAGCATCATCACCACCAAG 217
DB 156 ASPALAGLIHNSIALLALEUASPRVALALEUAGIUSERILEGLYVALSPHTRASN 175

```

```

QY 218 GTTTCGCCAAGAGAGAAATCTGCCGCGACACATGATCAGCTTTGTGATCCTT 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 176 LeuLeuVal---SerGlnProAspGlyAlaGlnAlaLeuSerLeuValAspThrLeu 194
QY 278 GTGACCAATTACACAGTGTGACTTCAGAAATAGAGCACTGCTGTGATTGAGATATA 337
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 195 IleArgSerGlySerVal-----AspVal 202
QY 338 TTGGCCGAGCAAAATTCCTATTGTGTGGAGGAGACCAATTATTACATTGATCTCG 397
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 203 ValValValAspSerValAlaAlaLeuVal----- 212
QY 398 CTCTGAAAGTCTGTGATACCAAGCCCGAGAGATGGCAGCTGAGAAAGTATTGAC 457
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 212 ----- 212
QY 458 CGAAAGTGGACCTTGAAAGAGAGATGCTTGTACTTCACAAACGCGCTAAGCCAGTG 517
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 213 ProLysThrGluLeuAspLysLysLysLysLysLysLysLysLysLysLysLysLys 222
QY 518 GACCCAGAAATGCT-----GCCAAGCTGCATCCACATGACAAACGCAAGTGGCCAGG 571
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 223 AspAlaIleValAlaLeuGlnAlaLeuMetSerGlnAlaLeuArgLysLeuSerHis 242
QY 572 AGCTTGCAAGTTTGAAGAAAGAAATCTCTCATAGTGAATTTCTCCATGCTCAACT 631
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 243 SerLeuSerLeuSerGlnThrValLeuLeuPheIleAsnGlnIleArgAlaLysValAla 262
QY 632 ACGAAGAGTGTGTGTCCTT-----GGAGGTCCTGGAAGTTCTCT 676
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 263 ThrPheGlyPheGlyGlyProThrGlnValThrSerGlyGlyAsnAlaLeuLysPhe 281
QY 677 AACCTTGACATCTTGTGCTCATGCTGACGAGCATTTCTAGATGAGCGCTTG----- 730
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 282 -----TyrAlaSerValArgLeuAsnIleArgArgIleGlyPhe 294
QY 731 -----GATTAAGAGGTGATGACATGCTTGTGCTGCGCTCTTGAG----- 772
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 295 LeuLysGlnGlyGlnGlnIleThrIleGlySerGlnValAlaValLysLysLysLys 314
QY 773 -----GAACTAAGAGATTTTTCACAGACGCTATATACAGAGATGTTTGC 817
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 315 HisAlaProProPheLysThrAlaGlnPheGlnLeuGlnPheGly---LysGlyLys 333
QY 818 GAAATAGCCAGACTATACATGCTATC----- 847
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 334 ArgSerSerGlnLeuPheGlnLeuGlyLeuLysHisLysLysLysLysLysLysLys 353
QY 848 -----TCCCAATCAATGCTGCTCAG-----GAATTTGACGAGTACCTG 886
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 AlaTyrTyrArgPheAsnAspMetSerPheLysGlyLysAsnAsnLeuLysSerTyrLeu 373
QY 887 ATCACTAGAGAAATGACACAGTGAAGTAAACCACTTAAAGAAAGAGCTGCT 946
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 374 ---ThrGlnAsnLysSerValAlaAsnAspLeuGlnIleThrLysLeuArgArgLeuMetGly 392
QY 947 CCCATTGCCCCCTGCTATGCTTGAAGGTATGATGTTCCAGAGTGGAGAGCTCT 1006
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 393 ThrGlnAlaProLysGlnGlnIleThrLysLysSerSerProSerAspLeuProGlnGln 412
QY 1007 GTTCTTGAACT 1018
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 413 ValValThrPro 416

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 587966440 No. 58796644disk of No. 58796644th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557,122A
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambirth, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3980.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 3052 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-557-122A-26

Alignment Scores:
Pred. No.: 0.465 Length: 3052
Score: 95.50 Matches: 86
Percent Similarity: 36.91% Conservative: 62
Best Local Similarity: 21.45% Mismatches: 177
Query Match: 2.67% Indels: 76
DB: Gaps: 16

US-09-513-151-3 (1-2041) x US-08-557-122A-26 (1-3052)
QY 10 GATGGCGTGGCGGCGGCGACAGACAGTCTGTGGGCGAGGCGCTCAGGGCCCTGCA 69
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1028 AspAspIleProPheGlyLeuThrAlaSerSerAspAspAlaAlaIleGlnThrLeu 1047
QY 70 ACGAACCCTACTCTTGTAGTATCTCGGGCCACAGGCGACCGGCAAAATCCACGCTGC 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1048 ValAspSerSerGlnValValIleGlyPhePheLysAspValThrSerAspAlaAl 1067
QY 130 GTTGCAGCTAGGCGACGCGCTCGCGGTAGATCGTC----- 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1067 AlysGlnPhe-----LeuLeuAlaIleGlnSerValAspAspIleProPheGlyLys 1085
QY 167 -----AGCGTGATCCATCAGGTCTATGAGCGCTAGACATCACCACCAAGGTTTC 222
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1085 rSerSerAlaAspAspLeuProAlaTyrLeuAlaAsnGlnIleThrPheValIleProValIle 1105
QY 223 TGCCCAAGAGAGAGATCTGCCGCGACCAACATGATCAGCTTGTGTGATCTCTGTGAC 282
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1105 eValGlnSerGlyLysIleAspAlaAspPheAsnAlaThrPheTyrSerMetAlaAsnLys 1125
QY 283 CAATTACACAGTGTGAGATTCAAGAAATAGACCAACTGCTCTGATTGAAGATATATTGC 342
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1125 sHisPheAsnAspTyrAspPheValSerIleGlnLysAla-----AspValPheSe 1142
QY 343 CCGA-----GACAAATTCCTATTGTTG-----GGAGGAGC 375
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1142 rLysTyrGlnLeuAspLysAspGlyValValLeuPheLysLysPheAsnGlnGlyArgAs 1162
QY 376 CAATTATTACATGATCTCTGCTGTAAGTCTTGTGCAATTACCAAGCCCGAGAGAT 435
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1162 nasnPhGlnGlyGlnIleThrLysGlnLysLeuLeuAspPheIleLysHisAsnGlnIle 1182

```

```

QY 436 GGGCACTGAGAAAGTATTCACGAAAGTGACCTTGAAAGAGATGCTCTTGACT 495
      :          :          :          :          :          :
Db 1182 uProleu---ValIleValPheSerIysTyrGlnLeuAspIys---AspGlyValIleVal 1200
QY 496 TCACAAACGGCTA-----AGCCAGGTGGACCCCAAGAAATGGC 531
      :          :          :          :          :          :
Db 1200 uPheLysLysPheAspGlnGlyArgAsnAsnPheGlnGlyValThrLysGlnLysLe 1220
QY 532 TGCCAAAGCTGCATCCATGACAAACGCAAGGAGGAGGCTTGCAATTTTGAAGA 591
      :          :          :          :          :          :
Db 1220 uLeuAspPheIleLysHisAsnGlnLeuProLeu-----ValIleValPheSerLys 1237
QY 592 AACCAAGATCTCTCATAGTGAATTTCTCCATCTGCACATACGGAAGAGGTGGTGC 651
      :          :          :          :          :          :
Db 1237 sTyrGlnLeuAspLysAspGlyValIleLeuPheLysLysPheAspGlnGlyArgAsnAs 1257
QY 652 CCTTGAGAGTCTCTGAAGTCTCTAAC-----CC 681
      :          :          :          :          :          :
Db 1257 nPheGlnGlyGlnValThrLysGlnAsnLeuLeuAspPheIleLysHisAsnGlnLeuPr 1277
QY 682 TTGCATCTCT-----TGACCTTCATGCTGACGACGAGTCTAGATGACGC 726
      :          :          :          :          :          :
Db 1277 oLeuValIleValPheSerArgTyrGlnValHisGlnAspGlyValIleLeuPheLysLys 1297
QY 727 CTTCGATPAGAGGCTGATGACATGCTGCTGCTGGCTCTTGGAAGACTPAGAGATT 786
      :          :          :          :          :          :
Db 1297 sPheAspGlnGlyArgAsnAsnPheGlnGlyValThrLysGlnLysLeuAspPhe 1317
QY 787 T-----CACAGAGCTATATCAGAAATGTTGCGAAATACCAGGAGCTATACACA 840
      :          :          :          :          :          :
Db 1317 eIleLysHisAsnGlnLeuProLeuValIleValPheSerLysTyrGlnLeuSerGlns 1337
QY 841 TGTATCTTCCATCAATTCGCTTCAGAGAAATTCACGATACCTGATCACT--GAGAG 897
      :          :          :          :          :          :
Db 1337 pGlyVal-----ValLeuPheLysLysPheAspGlnGlyArgAsnAsnPheGlnG 1354
QY 898 AAAATGCACACTGAGACTACTAACACACCTTCTAAGAAAGACCTGCTCCATTGTC-- 955
      :          :          :          :          :          :
Db 1354 yAspLeuThrLysAspAsnLeuLeuAsnPheIleLysSerAsnGlnLeuProLeuValI 1374
QY 956 -----CCCCCGCTGCTATGG 969
      :          :          :          :          :          :
Db 1374 eAspAspPheLysLeuSerIleTyrLeuProSerIleMetAspGlnProValIleTyrAs 1394
QY 970 CTTAGAGGTATCTGATGCTGCGAAGTGGAGAGACTGCTTCTTGACCTGCTTGAAT 1029
      :          :          :          :          :          :
Db 1394 nGlyLysLysAlaAspIleAla-----AspAlaAspValPheGlnLysTyrPheGlnI 1412
QY 1030 C 1030
Db 1412 1 1412

```

```

RESULT 12
US-09-262-666-26
: Sequence 26, Application US/09262666
: Patent No. 6346244
: GENERAL INFORMATION:
: APPLICANT: Hjort, Carsten Malland
: TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
: NUMBER OF SEQUENCES: 38
: CORRESPONDENCE ADDRESS:
: ADDRESSSEE: No. 63462440 No. 63462444 disk of No. 63462444th America, Inc.
: STREET: 405 Lexington Avenue, 64th floor
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10174-6401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/262,666
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/557,122
: FILING DATE: 11-DEC-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Lambiris, Elias J.
: REGISTRATION NUMBER: 33,728
: REFERENCE/DOCKET NUMBER: 3980, 204-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-0123
: TELEFAX: 212-878-9655
: INFORMATION FOR SEQ ID NO: 26:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3052 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-09-262-666-26

Alignment Scores:
Pred. No.: 0.465 Length: 3052
Score: 95.50 Matches: 86
Percent Similarity: 36.91% Conservative: 62
Best Local Similarity: 21.45% Mismatches: 177
Query Match: 2.67% Indels: 76
Gaps: 16

US-09-513-151-3 (1-2041) x US-09-262-666-26 (1-3052)
QY 10 GATGCGCTCCGTCGCGGCTGACAGAGCAAGTTCCTGGGCACTGAGGCTCAGGGCGCTGCA 69
      :          :          :          :          :          :
Db 1028 AspAspIleProPheGlnGlyLeuThrAlaSerSerAspAlaAlaIleValThrLeu 1047
QY 70 ACGGACCTTACTCTTGATGATCTCGGGCCACGGGACCGGCAATCCACCTGGC 129
      :          :          :          :          :          :
Db 1048 ValAspSerSerGln-ValIleValIleGlyPhePheLysAspValThrSerAspAla 1067
QY 130 GTTGAGCTAGGCGGCGGCTGCGGCGGAGATCGTC----- 166
      :          :          :          :          :          :
Db 1067 aLysGlnPhe-----LeuLeuAlaIleAlaGlnSerValAspAspIleProPheGlyI 1085
QY 167 -----AGCGTGCATCTCATGCTGATGAGAGGCTTAGACATACACCAAGTTTC 222
      :          :          :          :          :          :
Db 1085 rSerSerAlaAspAspLeuProAlaTyrLeuAlaAsnGlnThrPheValThrProValI 1105
QY 223 TGCCCAAGAGCAGAGAAATCTGCCGCGCACCATGATCAGCTTTGTGATCCTCTTGTC 282
      :          :          :          :          :          :
Db 1105 eValGlnSerGlyLysIleAspAlaAspPheAsnAlaThrPheTyrSerMetAlaAsn 1125
QY 283 CAATTACACAGTGTGGAGTCTGCAAGAAATGAGCAACCTGCTGATTGAATATATTGC 342
      :          :          :          :          :          :
Db 1125 sHisPheAsnAspTyrAspPheValSerIleGlnAlaAsnAla-----AspValPhe 1142
QY 343 CCGA-----GACAAATTCCTATTGTTGTG-----GGAGGAAC 375
      :          :          :          :          :          :
Db 1142 rLysTyrGlnLeuAspLysAspGlyValValLeuPheLysLysPheAspGlnGlyArgAs 1162
QY 376 CAATTATTAATTAATCTCTGCTGCGAAGCTTTGTCAATACCAAGCCCAAGAGAT 435
      :          :          :          :          :          :
Db 1162 nAsnPheGlnGlyGlnIleThrLysGlnLysLeuLeuAspPheIleLysHisAsnGln 1182
QY 436 GGGCACTGAGAAAGTATTCACGAAAGTGACCTTGAAAGAGATGCTCTTGACT 495
      :          :          :          :          :          :
Db 1182 uProleu---ValIleValPheSerLysTyrGlnLeuAspIys---AspGlyValIle 1200
QY 496 TCACAAACGGCTA-----AGCCAGGTGGACCCCAAGAAATGGC 531
      :          :          :          :          :          :
Db 1200 uPheLysLysPheAspGlnGlyArgAsnAsnPheGlnGlyValThrLysGlnLysLe 1220

```

```

OY 532 TGCGAAGCTCATCATGACGAAGGCGGAGAGCTTGCAAGTTTGTGAGA 591
      |||:||||:
Db 1220 ulcuasppheilelyshisnslneuprou-----ValilealpheSerly 1237
OY 592 AACGGAATCTCATAGTGAATTTCTCCATCGTCAACATACGAGAGAGTGGTGC 651
      |||:||||:
Db 1237 stYrGlnleuAspPlyAspGlyValValleuPheLysPheAspGlnLysArgAsnAs 1257
OY 652 CCTTGAGAGTCTCTGAGAGTCTCTTAC-----CC 661
      |||:||||:
Db 1257 nphesgluglyValThrlYsgLysnleuAspPheilelyshisnslneupr 1277
OY 682 TTGCATCCTT-----TGCTTCATGCTGACGAGAGCTTGATAGAGCG 726
      |||:||||:
Db 1277 oleuValilealpheSerArgTyrGlnValHisGlnAspGlyValValleuPheLysly 1297
OY 727 CTTCGATTAAGAGGTGATGATGCTTGGCTGCTGAGAGTGAAGTGAAGATTT 786
      |||:||||:
Db 1297 sphaespglyLysArgAsnAsnphesgluglyValThrlYsgLysleuAspPh 1317
OY 787 T-----CACAGACGCTATATACGAAGAATGTTGGAAATAGCCAGACTATCAACA 840
      |||:||||:
Db 1317 eilelyshisnslneuprouValilealpheSerLysTyrGlnleuSerGlnAs 1337
OY 841 TGCTATCTTCATCAATGCTTCAAGAGATTTCAAGAGTACCTGATCACT---GAGGG 897
      |||:||||:
Db 1337 pglYVal-----ValleuPheLysPheAspGlyLysArgAsnAsnphesglugly 1354
OY 898 AAAATCAGACGTGAGAGCTAGTACAGAGCTTCTAAAGAAAGAGCTGGTCCATTGTC- 955
      |||:||||:
Db 1354 yAspLeuThrlYsAspAsnleuAsnphellelySerAsnGlnleuProuValil 1374
OY 956 -----CCCCCTGTCATGCG 969
      |||:||||:
Db 1374 eaAspAspPheLysleuSerileTyrleuProuSerAlaMetAspGlnProuValYlYras 1394
OY 970 CTAGAGATTCATGATGCTCGAAGTGGAGAGTCTGCTTGAACCTCTCTGGAAT 1029
      |||:||||:
Db 1394 nglYsLysAlaAspPhelela-----AspAlaAspValPheGlnLysTrpLeuGlnVa 1412
OY 1030 C 1030
      |||:||||:
Db 1412 1 1412

RESULT 13
Sequence 3, Application US/0836442
GENERAL INFORMATION:
APPLICANT: DOCHERTY, Andrew, J.P.
APPLICANT: SLOCUMBE, Patrick, M.
TITLE OF INVENTION: A HUMAN METALLOPROTEINASE
TITLE OF INVENTION: VARIANTS THEREOF AND DNA SEQUENCES CODING THEREFOR
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,442
FILING DATE: 01-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/02181
FILING DATE: 13-MAR-1997

```

```

APPLICATION NUMBER: GB 9612150.4
FILING DATE: 11-JUN-1996
APPLICATION NUMBER: GB 9526229.1
FILING DATE: 21-DEC-1995
APPLICATION NUMBER: GB 9521498.7
FILING DATE: 20-OCT-1995
APPLICATION NUMBER: GB 95521495.3
FILING DATE: 20-OCT-1995
APPLICATION NUMBER: GB 9518023.8
FILING DATE: 05-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 47425
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 529 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 5990293e
US-08-836-442-3

Alignment Scores:
Pred. No.: 0.213 Length: 529
Score: 94.00 Matches: 72
Percent Similarity: 32.93% Conservative: 63
Best Local Similarity: 17.56% Mismatches: 161
Query Match: 2.638 Indels: 114
DB: Gaps: 19

US-09-513-151-3 (1-2041) x US-08-836-442-3 (1-529)
OY 413 GTCAATACCAAGCCGAGAGGCGACATGAAAGTATGACGGAAGTGAGCTT 472
      |||:||||:
Db 95 lIeThrlThSerProGlnIlelel----- 102
OY 473 GAAAGAGAGAT-----GGTCTTGTACTTCACAAAGCCCTAACGAGTGAC 520
      |||:||||:
Db 103 -----AspAspCysTyrTyrGlnGlnYHisileuAsnGlnLysValIserAspAlaSer 120
OY 521 CCAGAAATGCTGCCAAGCTGCATCCACAT-----GACAAAGCCAAAGTGCGC 568
      |||:||||:
Db 121 lIeSerThrlCysArgGlyLeuArgGlyTyrPheSerGlnGlyAspGlnArgTyrPheile 140
OY 569 AGGAGCTGCAAGATTTTGAAGAAAGCAATCTCATAGTGAATTTCTCCATGTCGCA 628
      |||:||||:
Db 141 GluProuSerProIleHisArgAspGlyGlnGlnHisAlaPheLysTyrAsnPro 160
OY 629 CATACGGAAGAGGTGGTGGTCCCTTGAGGCTCTGTAAGTCTCTAACCTTGC-- 685
      |||:||||:
Db 161 AspGlnLys-----AsnTyrAspSerThrlCysGly 170
OY 686 -----ATCCTTGCGCTTCATGCTGACGACGAGCACTTCAATGAGCGCTTGATAG 736
      |||:||||:
Db 171 MetAspGlyValleuTrpAlaHis-----AspLeu 180
OY 737 AGGAGTATGATGATGCTGCTGCTGAGAGTCTTGAGAGTGAAGAT----- 784
      |||:||||:
Db 181 GlnGlnAsnIleAlaProuAlaThrlYsleuValLysleuLysAspArgLysValGln 200
OY 785 TTTCAGAGAGCTATATACAG-----AAGAAATGTT 814
      |||:||||:
Db 201 GlnHisGlnLysTyrIleGlnTyrTyrleuValleuAspAsnGlnGlyLysPheLysArgTyr 220
OY 815 TCGGAATATAGCCAGACTATCAACATGATGATCTTCCA-----TCAATT 839
      |||:||||:
Db 221 AsnGlnAsnGlnAspGlnIleArgLysArgValPheGlnMetAlaAsnTyrValAsnMet 240

```

```

OY 860 GGCCTCAAGAAATTCACAGTACCTGATCACTAGAGAAATGACACTGAGACTAGT 919
    ::::: ::::: :::::
Db 241 LeuTyrLysLysLeuAsnThrHisValAlaLeuValGly---MetLuleIerPrThrAsp 259
OY 920 AACGACCTTCAAGAAAGAGACCTGCTCCATTTGCCCTGTCTATGAGGTA 979
    :::::
Db 260 LysAspLysIleLys-----IleThrProAsnLaseSerPheThrLeu 273
OY 980 TGTGATGTCGCAAGGAGAGAGTCTGTTCTTGAA----- 1015
    :::::
Db 274 GluAsnPheSerLysThrArgLysValLeuSerArgArgLysArgHisAspIleAla 293
OY 1016 -----CTGCTCTTCAAAATCGTCAAGT----- 1039
    :::::
Db 294 GluLeuIleThrAlaThrGluLeuAlaGlyThrValGlyLeuAlaPheMetSerThr 313
OY 1040 -----TTTCATCAGAGGCCACAGACCTGACCCATCCATCAATA 1075
    :::::
Db 314 MetCysSerProTyrSerValGlyValAlaGlnAspHisSerAspAsnLeuLeuArgVal 333
OY 1076 -----AGATCCATACATGACAGCTGAGAACAGAGATTATCAC-----CTG 1120
    :::::
Db 334 AlaGlyThrMetAlaHisGluMetGlyHisAsnPheGlyMetPheHisAspTyrSer 353
OY 1121 TGTGACCTGTGATGCAATCATCATGCGGATCGGATCGGACCGACCGACATCAATCC 1180
    :::::
Db 354 CysLysCysProSerThrIleCysValMetAspLysAlaLeuSerPheTyrIleProThr 373
OY 1181 AAA---TCCCACTTGAAACCACTGAGAAAGAAAGAAAGATTGACATGCTGTCAAC 1237
    :::::
Db 374 AspPheSerSerCysSerArgLeuSerTyrAspLysPhePheGluAspLysLeuSerAsn 393
OY 1238 ACCATGAGAAAGTCAGAGCTTTCCCGACATAACAAAGAACTTAAGGGAAGGATCC 1297
    :::::
Db 394 CysLeuPheAsnAlaProLeuProThrAspIleIleSerThrProIleCysGlyAsnGln 413
OY 1298 CCAGGGCAAGATGATCAAGCTGAATGACAGGTTTAAAGACATGCTCCAGG----- 1351
    :::::
Db 414 LeuValGluMetGlyGluAspCysAspCysGlyThrSerGluThrCysLysIleLysAla 433
OY 1352 GCGTTTGGAAGAGTGTGGG-----ATCCAGTTTCAGAGGAGGAGGTAT 1396
    :::::
Db 434 ThrPheGlnCysAlaLeuIleGlyCysGlyLysGlnPheLysLysAlaGlyMet 453
OY 1397 GTTGTCTCCAGTGTGGCCAAAGAGTGC 1426
    :::::
Db 454 ValCysArgProAla---LysAspGluCys 462

```

RESULT 14

```

US-09-040-725A-1
; Sequence 1, Application US/09040725A
; Patent No. 639584
; GENERAL INFORMATION:
; APPLICANT: Institut Curie
; APPLICANT: CNRS
; APPLICANT: Arpin, Montique
; APPLICANT: Crepaldi, Tiziana
; APPLICANT: Gautreau, Alexis
; APPLICANT: Louvard, Daniel
; TITLE OF INVENTION: Pharmaceutical composition containing ezrin mutated
; FILE REFERENCE: 39108200100
; CURRENT APPLICATION NUMBER: US/09/040,725A
; CURRENT FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-040-725A-1

```

Alignment Scores:

```

Pred. No.: 0.228
Score: 94.00
Percent Similarity: 32.068
Best Local Similarity: 19.148
Query Match: 2.63%
DB: 4
Gaps: 16

US-09-513-151-3 (1-2041) x US-09-040-725A-1 (1-586)

OY 164 GTCAGCGGTGACCTCCATGACGCTATGAGGCTTAGACATATCCAAAGGTTTCT 223
    :::::
Db 220 ValAspAlaLeuGlyLeuAsnIleTyrGluLysAspAspLysLeuThrProLysIleGly 239
OY 224 GCCCAAGACAGAAATCTGCCGCCACCATGATGACTT----- 265
    :::::
Db 240 PheProTyrSerGluIle-----ArgAsnIleSerPheAsnAspLysPheVal 256
OY 266 -----GTGATCCTCTTGTGACCAATTCACAGCTGCTGACTTGCAATATAGACA 316
    :::::
Db 257 IleLysProIleAspLysLysAlaProAspPheValPheTyrAlaProArgLeuArgIle 276
OY 317 ACTGCTGTGATT-----GAAGATATATTGCCCCGAGCAAAATT 355
    :::::
Db 277 AsnLysArgIleLeuGlnLeuCysMetGlyAsnHisGluLeuTyrMetArgArgArgLys 296
OY 356 CCTATTGTTGTGGAGAACCAATTATTTCATTGAATCTGCTGCTGAAAGTTCTTGTG 415
    :::::
Db 297 ProAspThrIleGluValGlnGlnMetLysAlaGlnAlaArgGluLysHisGlnLys 316
OY 416 AATACCAAGCCCGAGAGATGGGCACTGAGAAAGTATGACCGAAATGAGCTTGA 475
    :::::
Db 317 GlnLeuGluArgGlnGlnLeuGlnIleGlnLysLysArgArgIleThrValGluArgGlu 336
OY 476 AAGGAGATGCTTCTTACTTCACAAACGCTTAAGCCAGTGGAGCCAGAAATGCTGCC 535
    :::::
Db 337 LysGlu-----GlnMetMetArgLysGluGluLeuMetLeu 349
OY 536 AACCTGCATCCATGACAAAGC-----AAAGTG 565
    :::::
Db 350 ArgLeuGlnAspTyrGluGluLysThrLysLysAlaGluArgLysLeuSerGluGlnIle 369
OY 566 GCGAGAGCTGCAAGTTTGTGAAG----- 592
    :::::
Db 370 GlnArgAlaLeuGlnLeuGluGluArgLysArgAlaGlnGluAlaGluArgLys 389
OY 593 -----ACAGAAATCTCTCATAGTAATTTCTCCATGCTCAACATACGAA 637
    :::::
Db 390 GluAlaAspArgMetAlaAlaLeuArgAlaLysGluGluLeuGlnArgGln 406
OY 638 GAAGTGTGTGTCCTTGAGAGTCTCTGAAGTCTTAACCTTGATCTTGGCTT 697
    :::::
Db 406 ----- 406
OY 698 CATGCTGACCGGCACTTCTAGATGAGCGCTTGATGAAGGCTGATGACATGCTTGT 757
    :::::
Db 407 AlaValAspGlnIleLysSerGlnGlnLeuAlaAlaGluLeuAlaGluTyrThrAla 426
OY 758 ---GCTGGGCTTTTGAGAACTAGAGATTTTTCACAGACGCTAATTCAGAAATGTT 814
    :::::
Db 427 LysIleAlaLeuLeuGlnGluAlaArg-----ArgArgLys 438
OY 815 TCGGAAATATACCGACGACATATCAACATGATCTTCAATCAATTGGCTTCAAGAAATT 874
    :::::
Db 439 GluAspGluValGlnGluTyrPheGlnHisArgAla-----LysGluAla 452
OY 875 CACGACTACCTGATCACTAGAGGAAATGACACACTGAGACTAGTACACAG-----CTT 928
    :::::
Db 453 GluAspAspLeuVal-----LysThrLysGluLeuLeuHisLeu 465
OY 929 CTAAGAAAGACCTGTCCTCATGTCCTGCTATGAGCTTATGAGGATATCTGATGTC 988
    :::::
Db 466 ValMetThrAlaProProProProProProProProProProProProProProPro 478

```

QY 989 TCGAGTGGGAGGAGTCTGTTCTTGAACCTGCTTGAATCGCAAGTTTCATCCAG 1048
 Db 479 -----GluProValSerThrHisValGlnSerLeuGln 490
 QY 1049 -----GGCCACAGCCTACAGCCACTCCATAAAGATGCCATACATGAAGCTAGAAC 1102
 Db 491 AspGluGlyAlaGluProThrGly----- 498
 QY 1103 AAGAGAGTTATCACCCTGTGACCTGTGTGATCGAATCATCATTTGGGATCGCAATGG 1162
 Db 499 -----TyrSerAlaGluLeuSerSerGluGlyIleArgAspAspArg----- 512
 QY 1163 GCAGGCGACATTAATCCAAATCCCACTTGGAACCACTGAGAAAGAAAGAAATGGAC 1222
 Db 513 -----AsnGluGluLysArgIleThrGluAlaGluLysAsnGluArgValGln 528
 QY 1223 TCAGATGCTGCACACACATAGAAAGTGCAGAGCTGTTCCCGCAGATATACAA 1276
 Db 529 ArgGlnLeuValThrLeuSerSerGluLeuSerGlnAlaArgAspGlnAsnLys 546

RESULT 15

US-08-836-943-2
 ; Sequence 2, Application US/08836943
 ; Patent No. 5965391

GENERAL INFORMATION:

APPLICANT: Reinscheid, Dieter
 APPLICANT: Elkmanns, Bernhard
 APPLICANT: Sahm, Hermann
 TITLE OF INVENTION: DNA WHICH REGULATES GENE EXPRESSION IN
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: The Firm of Karl F. Ross, PC
 STREET: 5676 Riverdale Ave.
 CITY: Bronx
 STATE: New York
 COUNTRY: USA
 ZIP: 10471

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/836,943
 FILING DATE: 08-MAY-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Myers, Jonathan
 REGISTRATION NUMBER: 26,963
 REFERENCE/DOCKET NUMBER: 20357
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (718) 884-6600
 TELEFAX: 718/601-1099
 TELEX: 620428

SEQUENCE CHARACTERISTICS:
 LENGTH: 739 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-836-943-2

Alignment Scores:
 Pred. No.: 0.266 Length: 739
 Score: 94.00 Matches: 70
 Percent Similarity: 37.24% Conservatives: 57
 Best Local Similarity: 20.53% Mismatches: 136
 Query Match: 2.63% Indels: 78
 DB: 2 Gaps: 15

US-09-513-151-3 (1-2041) x US-08-836-943-2 (1-739)

QY 89 GTGATTCGCGGGCCACGGGCACC-----GGCAATCCACGCTGGCGTTG 133
 Db 336 ValTyrIleGlyArgAsnGlyThrGluLeuValLeuHisGlyArgSerLeuPheVal 355
 QY 134 CAG---CTAGGCGACGGCTGGCGGTAGATTCGACCGGCTGACTCATGAGGTAT 190
 Db 336 ArgAsnValGlyHisLeuMetGlnAsnProSerIleLeuAspGlyGluGlnIlePhe 375
 QY 191 GAAGGC---CTAGCATCATCCACCAACAGGTTTCTGCC----- 226
 Db 376 GluGlyIleMetAspAlaValLeuThrThrValCysAlaIleProGlyIleAlaProGln 395
 QY 227 CAAGAGCAGAAATCTGCCGACCATGATCATGCTTTGGATCCCTTGTGACCAAT 286
 Db 396 AsnLysMetArgAsnSerArgLysGlySerIleTyrIleValLysProLysGlnHisGly 415
 QY 287 TACACAGTGGGAGCTTACAGAAATGACCACTGCTGTGATGAAGATATATTTGGCCGA 346
 Db 416 ProGluGluValAlaAlaPheThrAsnGluLeuPheGlyArgValGlnAspLeuAspLeu 435
 QY 347 GACAAATTCCTATTGTTGGGA-----GGAAACCAATTATTACATT 388
 Db 436 ProArgHisThrLeuLysValGlyValMetAspGluGluArgThrSerValAsnLeu 455
 QY 389 GAATCTGCTCTGGAAGTT-----CTTGCAATACCAAGCCCGAGAG 433
 Db 456 AspAlaSerIleMetGluValAlaAspArgLeuAlaPheIleAsnThr----- 471
 QY 454 ATGGGACTGAGAAAGTATGACCGA-----AAAGGAGCACTGAAAGAGAGAT 484
 Db 472 -----GlyPheLeuAspArgThrGlyAspGluIleHisThrSerMetGluAla 487
 QY 485 GGTCTTGACTTCACAAAGCGCTAAGCCAGTGAGCCAGAAATGGCTCCAGCTGAT 544
 Db 488 GlyAlaMetValArgLysAlaAspMetGlnThrAlaProTyrGlnAlaTyrGlnAsn 507
 QY 545 CCACATGACAAAGCGAAAGTGCGCCAGAGCTTGCAAGTTTGAAGAAGCAAGATCTCT 604
 Db 508 AsnAsnValAspAlaGlyIleGlnArgGlyLeuProGlyLysAlaGlnIleLeuLysGly 527
 QY 605 -----CATAGTAATTTCCCATCGCAACATACAGAAAGAGT 643
 Db 528 MetThrAlaMetThrGluLeuMetValGlnMetLeuGluLysLys----- 542
 QY 644 GGTGTCCTCCCTTGAGGTCCTCTGAAGTTCTTAAC-----CCTTGC 685
 Db 543 -----IleGlyGlnProArgGluGlyAlaAsnThrAlaTyrValProSerProThr 559
 QY 686 ATCCTTGGCTTCATCGACGACGACGACGATCTAGATGAGCGCTTGGAATAGAGGTGAT 745
 Db 560 GlyAlaThrLeuHisAlaThrHisTyrHisLeuValAsp---ValPheLysValGlnAsp 578
 QY 746 GACATGCTGCTGCTGGGCGCTCTTGAGGAGACTAGAA----- 781
 Db 579 GluLeuArgAlaAlaGlyArgArgAspSerLeuArgAsnIleLeuThrIleProThrAla 598
 QY 782 -----GATTTTCACAGACGCTATATATCAGAAATGTTTCGAAATAGCCAGAGAC 832
 Db 599 ProAsnThrAsnTyrSerGluGluGluLysLysGluMetAspAsnAsnCysGlnSer 618
 QY 833 -----TATCAACATGGTATCTTCATCAATCATATGGCTTCAAG 868
 Db 619 IleLeuGlyTyrValValArgTyrValGlnHisGlyValGlyCysSer---LysValPro 637
 QY 869 GAATTCACGAGTACCTCATCATGACGAGGAATCCACACAGGAGACCTAGTACCAAGTT 928
 Db 638 AspIleHisAspIleAspLeuMetGlnAspArgAlaThrLeuArgGlyIleSerSerGlnMet 657
 QY 929 CTA 931
 Db 658 Leu 658

Tue Apr 22 09:11:09 2003

us-09-513-151-3.n2p.rai

Page 15

Search completed: April 21, 2003, 18:52:17
Job time : 42.0286 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 21, 2003, 17:14:14 ; Search time 24.423 Seconds
(without alignments)
6932.240 Million cell updates/sec

Title: US-09-513-151-3
Perfect score: 3575
Sequence: 1 CTGCCATTAAGATGCGCTCCG.....TTTACACAGAAAAAAAAAAAA 2041

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+.n2p.model -DEV=xlh
-O=/cgn2.1/USPPO.spool/US09513151/runat.15042003.141143.26366/app.query.fasta_1.2446
-DB=SwissProt_40 -QFMT=fasta -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPT=0
-LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human.0.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09513151_cgn2.1.1.25 @runat.15042003.141143.26366 -ICPU=3
-NO_XIPXY -NO_MMAP -LARGEOVER -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	548	15.3	428	MOD5_YEAST	P07884 saccharomyc
2	417	11.7	314	MIAA_BACHD	O9KAC3 bacillus ha
3	403.5	11.3	314	MIAA_BACSU	O8Y713 bacillus su
4	397.5	11.1	305	MIAA_LISMO	O8Y713 listeria mo
5	395.5	11.1	305	MIAA_LISIN	O92C59 listeria in
6	378.5	10.6	309	MIAA_CLOAB	O97121 clostridium
7	377.5	10.6	310	MIAA_CLOPE	O8X185 clostridium
8	363	10.2	323	MIAA_RALSO	O8XW60 ralsconia s
9	349	9.8	322	MIAA_PSEPU	O30762 pseudomonas
10	348	9.7	305	MIAA_THEMA	O9WV25 thermotoga
11	343.5	9.6	314	MIAA_CHLTR	O84771 chlamydia t
12	342.5	9.6	316	MIAA_PASMU	O9CM57 pasteurilla
13	339	9.5	311	MIAA_STEAM	O94490 staphylococ
14	334.5	9.4	311	MIAA_HAEIN	O94495 haemophilus
15	334	9.3	315	MIAA_VIBCH	O9KX12 vibrio chol
16	333	9.2	294	MIAA_LACIA	O9CHU2 lactococcus
17	329.5	9.2	295	MIAA_ANNSP	O8Y1N2 anababena sp
18	328.5	9.2	323	MIAA_PSEAE	O9H1U9 pseudomonas

19	327.5	9.2	314	1	MIAA_CHLNU	O9P1F7 chlamydia m
20	324.5	9.1	311	1	MIAA_MYCCE	P46811 mycobacteri
21	324.5	9.1	314	1	MIAA_MYCTU	O33232 mycobacteri
22	321.5	9.0	313	1	MIAA_NEIMA	O9JUN5 neisseria m
23	320	9.0	294	1	MIAA_STRPN	O971W5 streptococc
24	319.5	8.9	342	1	MIAA_CHLPN	O9Z6Z6 chlamydia p
25	318.5	8.9	299	1	MIAA_STRPY	O9A059 streptococc
26	318.5	8.9	313	1	MIAA_STRPE	O8Z1W3 versinia pe
27	314	8.8	305	1	MIAA_AQUAE	O67162 aquifex ae
28	312.5	8.7	317	1	MIAA_XYLEA	O9P1S6 xyella fas
29	311	8.7	316	1	MIAA_ECOLI	P16384 escherichia
30	310	8.7	316	1	MIAA_ECO57	O8XDN3 escherichia
31	305	8.5	312	1	MIAA_STRCO	O69967 streptomyce
32	302	8.4	316	1	MIAA_SALTI	O8Z1B6 salimonia
33	301.5	8.4	316	1	MIAA_BORBU	O51761 borrelia bu
34	301	8.4	316	1	MIAA_SALTY	P37724 salimonia
35	295.5	8.3	302	1	MIAA_HELPI	O9Z1J7 helicobacte
36	288	8.1	313	1	MIAA_RHIME	O92NR2 rhizobium m
37	287.5	8.0	311	1	MIAA_HELPI	O25961 helicobacte
38	286.5	8.0	306	1	MIAA_DEIRA	O9TR6 delinococcus
39	281.5	7.9	326	1	MIAA_RICCN	O92HW4 rickettsia
40	280	7.8	303	1	MIAA_STNY3	P74040 synecocyst
41	278.5	7.8	313	1	MIAA_NEMIB	O9J1R0 neisseria m
42	272	7.6	298	1	MIAA_AGR5	P38436 agrobacteri
43	265.5	7.5	321	1	MIAA_RHILU	O98K14 rhizobium l
44	259	7.2	315	1	MIAA_BUCAI	P57632 buchnera ap
45	256	7.2	295	1	MIAA_CAUCR	O9A6J5 caulobacter

ALIGNMENTS

RESULT 1
MOD5_YEAST STANDARD; PRT; 428 AA.
ID MOD5_YEAST
AC P07884; Q12203;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE tRNA isopentenyltransferase (EC 2.5.1.8) (Isopentenyl-diphosphate:
DE tRNA isopentenyltransferase) (IPTase) (IPTP).
GN MOD5 OR YOR274W.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=67172703; PubMed=3031457;
RA Najarian D., Dhanich M.E., Martin N.C., Hopper A.K.;
RT "DNA sequence and transcript mapping of MOD5: features of the 5'
RT region which suggest two translational starts.";
RL Mol. Cell. Biol. 7:185-191(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=8288C;
RA Cheret G., Bernardi A., Sor F.J.;
RT "DNA sequence analysis of the VP1-SNF2 region on chromosome XV of
RT Saccharomyces cerevisiae.";
RL Yeast 12:1059-1064(1996).
RN [3]
RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.
RX MEDLINE=92052176; PubMed=1946403;
RA Slusher L.B., Gillman E.C., Martin N.C., Hopper A.K.;
RT "RNA leader length and initiation codon context determine
RT alternative AUG selection for the yeast gene MOD5.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9789-9793(1991).
RN [4]
RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.
RX MEDLINE=91203856; PubMed=1850093;
RA Gillman E.C., Slusher L.B., Martin N.C., Hopper A.K.;
RT "MOD5 translation initiation sites determine N6-Isopentenyladenosine
RT modification of mitochondrial and cytoplasmic tRNA.";

RL Mol. Cell. Biol. 11:2382-2390(1991).

RN [5]

RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.

RX MEDLINE-94187700; PubMed-8139355;

RA Boguta M., Hunter L.A., Shen W.C., Gillman E.C., Martin N.C.,

RA Hopper A.K.;

RT *Subcellular locations of MOD5 proteins: mapping of sequences

RT sufficient for targeting to mitochondria and demonstration that

RT mitochondrial and nuclear isoforms commingle in the cytosol.*;

RL Mol. Cell. Biol. 14:2298-2306(1994).

CC -1- FUNCTION: RESPONSIBLE FOR THE MODIFICATION OF A37 TO ISOPENTENYL

CC A37 OF BOTH CYTOSOLIC AND MITOCHONDRIAL TRNAs.

CC -1- CATALYTIC ACTIVITY: Isopeentenyl diphosphate + tRNA = diphosphate +

CC tRNA containing 6-isopentenyladenosine.

CC -1- PATHWAY: BIOSYNTHESIS OF THE MODIFIED BASE ISOPENTENYLADEOSINE

CC IN TRNAs.

CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL, CYTOPLASMIC AND NUCLEAR.

CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS MAY BE PRODUCED FROM THE USE OF

CC ALTERNATIVE INITIATION CODONS IN THE SAME READING FRAME.

CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M15991; AAA34785.1;

DR EMBL: X89633; CAA61780.1;

DR EMBL: 275182; CAA99499.1;

DR PIR: A26711; A26717.

DR SGD: S0005800; MOD5.

DR InterPro: IPR002627; IPPF.

DR Pfam: PF01715; IPPF; 1.

DR ProDom: PD004674; IPPF; 1.

DR Trifams: TRF00174; miaa; 1.

KW Transferase; tRNA processing; ATP-binding; Alternative initiation;

KW Mitochondrion; Nuclear protein.

FT CHAIN 1 428

FT TRNA ISOPENTENYLTRANSFERASE, CYTOPLASMIC/

FT MITOCHONDRIAL ISOFORM.

FT TRNA ISOPENTENYLTRANSFERASE, CYTOPLASMIC/

FT NUCLEAR ISOFORM.

FT FOR CYTOPLASMIC/NUCLEAR ISOFORM.

FT DOMAIN 210 232

FT BINDS ISOPENTENYLPYROPHOSPHATE

FT TRANSFERASE.

FT CONFLICT 313 313 MISSING (IN REF. 1).

FT CONFLICT 375 375 C -> R (IN REF. 1).

SO SEQUENCE 428 AA; 50236 MW; A956B17ABC05161F CRC64;

Alignment Scores:

Pred. No.: 6,94e-35 Length: 428

Score: 548.00 Matches: 143

Percent Similarity: 52.58% Conservatve: 81

Best Local Similarity: 33.57% Mismatches: 144

Query Match: 15.33% Indels: 58

DB: 1 Gaps: 15

US-09-513-151-3 (1-2041) x MOD5_YEAST (1-428)

QY 56 CTCAGGGC---CTGCAACGGACCTACTCTTGTAGTATCTCGGGCCACGGCCACC 112

DB 6 LeuysgLyCysLeuasmMetserLysValIleValIleAlaGlyThrThrGlyVal 25

QY 113 GGCAAAATCCACGCTGCGGTTCAGTAGGCGCAGCGCTCGCGGTTCAGTTCAGCGCT 172

DB 26 GlyLysSerGlnLeuSerIleGlnLeuAlaGlnLysPheAsnGlyGluValIleAsnSer 45

QY 173 GACTCCATGACGAGTGTATGACGAGCCATGACATCATCCACACAGGTTTCTGCCAAGG 232

DB 46 AspSerMetGlnValLysLysAspIleProIleIleThrAsnLysHisProLeuGlnGlu 65

QY 233 CAGAGATCTGCCGGCAGACATGATCAGCTTTGTGATCTCTTGTGACCAATTACACA 292

DB 66 ArgGlnGlyIleProHisIleValMetAsnHisValAsp---TrpSerGlnLysIleTyr 84

QY 293 GTGGTGCATTCAGAAATAGACCAACTGCTGATGATGATATATATATATATATATAT 352

DB 85 SerHisArgPheGlnThrGlnCysMetAsnAlaIleGlnAspIleHisArgArgGlyLys 104

QY 353 ATTCTATTGTTGTGGGAGAACCAATTATTCATTTGATGATCTGCTGTCGAAATTCCT 412

DB 105 IleProIleValIleGlyIleThrHisIleTyrLysLeuGlnThrLeuPheAsnLys---Arg 123

QY 413 GTCAATCCACAGCCCGCAGAGATGGC---ACTGAAAGTATGATGACCGAAAGTGAG 469

DB 124 ValAspThrLysSerSerIleArgLysLeuThrArgLysGlnLeuAsp-----Ile 140

QY 470 CTGAAAGAGAGATGGTGTGATCTACTTCACAAAGCCTTAAGCCAGGTGAGCCAGAAATG 529

DB 141 LeuGlnSerThrAspProAspValIleTyrAsnThrLeuValLysCysAspProAspIle 160

QY 530 GCTGCCAAGCTGCATCCATGACATGACAAAGCGAAAGTGCCAGAGCTTGCAGATTTTGA 589

DB 161 AlaThrLysTyrHisProAsnAspTyrArgArgValGlnArgMetLeuGlnLysIleTyr 180

QY 590 GAAACAGGAATCTCATGATGATTTCCATTCCTGACATACGAAAGAGAGTGTGT 649

DB 181 LysThrGlyLysLysProSerGlnThrPheAsnGlnLysIle----- 195

QY 650 CCCCTTGAGAGCTCTGTAAGTTCTCTACCTTCATCCCTTGATCGTTGATGATGACAG 709

DB 196 -----ThrLeuLysPhe---AspThrLeuPheLeuThrLeuIleTyrSerPro 210

QY 710 GCAGTTCTAGATGAGCGCTTGATGATGAGAGGTGATGATGATGATGATGATGATGATG 769

DB 211 GluProLeuPheGlnArgLeuAspArgValAspAspMetLeuGlnArgGlyAlaLeu 230

QY 770 GAGGAACTAGAGATTTTACAGACCGCTTATTCAGAAAGATTTTCCGAAATATACCGG 829

DB 231 GlnGlnIleLysGlnLeuLysGlnLysIleTyrSerGlnAsnLysPheThr-----ProGlu 248

QY 830 GACTATCAACATGATCTTCCATCAATCAATGATGCTTCAAGAAATTTTCAAGATGATC 889

DB 249 GlnCysGlnAsnGlyValIleProGlnValIleGlyPheLysGlnPheLeuProIleLeu 267

QY 890 ACTGAGGAAA-----TGC----- 904

DB 268 ---ThrGlyLysThrAspAspAsnThrValLysLeuGlnAspCysIleGlnArgMetLys 286

QY 905 -----ACACTGAGACAGTGTATCCAGCGCTTCRAAGAAAGACCTGCTCCATTTGCCCC 958

DB 287 ThrArgThrArgGlnTyrAlaLysArgGlnValLysTyrPheLysLysMetLeuIlePro 306

QY 959 CCTGTCTATGCG-----TTAGAGTATCTGATGATGTCGAGAGGAGAGAGTCT 1006

DB 307 AspIleLysGlnAspIleTyrLeuLeuAspAlaThrAspLeuSerClnThrAspThrAsn 326

QY 1007 GTTCTGAACCTGCTCTTGAATCGTCAAAAGTTTCATC-----CAG 1048

DB 327 AlaSerGlnArgAlaIleAlaIleSerAsnAspPheIleSerAsnArgProIleLysGln 346

QY 1049 GGCCACAGCCCTACAGCCCACTCCATMAAGATGCCATACATGAGAGCTGAGAACAGAA 1108

DB 347 GluArgAlaProLysAlaLeuGlnLysLeuLeuSerLysGlyLysThrThrMetLysLys 366

QY 1109 -----AGTTATCACCTGTGTGACCTCTCT-----GATGCA 1138

DB 367 LeuAspAspThrProThrHisTyrThrCysAsnValCysArgAsnAlaAspGlyLysAsnVal 386

QY 1139 ATCATCATTTGGGATGGCAATGGCAGCAGCCACATMAATCCCAATCCCACTTGAACAA 1198

DB 387 ValAlaIleGlyGlnLysTyrTrpLysIleHisLeuGlySerArgArgHisLysSerAsn 406

D	b	45	IleGlyThrAlaLysIleThrAlaGlUglUmetAspGlyValProHisIleLeuLeu	64
O	y	263	TTTGGAAGCTCCTTGTGACCAATTAAACAAGTGGTAGCTTCAGAAATAGCAACTGGCT	322
			: : : : : : : : : :	
D	b	65	IleLysAspPro-----SerGlnSerPheSerValAlaAspRheGlnAspLeuAlaThrPro	83
O	y	323	CTGATTAAGAATATAATTTGGCCCGAGACAAAATTCCTATTGTCTGGGAGGACCAATTAAT	382
			: : : : : : : : : : : : : : : : : : : :	
D	b	84	LeuIleThrGluIleHisGluArgGlyArgLeuProPheLeuValGlyGlyThrGlyLeu	103
O	y	383	TACATTAAGATCTCTGCCTTGGAAAGTT---CTGTCAATATCAAGACCCCAGAGATGGCG	439
			: : : : : : : : : : : : : : : : : : : :	
D	b	104	TyrValAsnAlaValIleHisGluPheAsnLeuGlyAspIleArgAlaAspGlu-----	121
O	y	440	ACTGAGAAAGTGAATGACCAGAAAGTGAGAGCTGA-----AAGAGATGGCTT	490
			: : : : : : : : : : : : : : :	
D	b	122	-----ASP TyrArgHisGluLeuGluValPheValAsnSerTyrGlyVal	136
O	y	491	--GTACTTCACAAAGCGCTTAAGCCAGATGGACCCAGAAATGGCTGCCAAGCTCATCCA	547
			: :	
D	b	137	GlnAlaLeuHisAspLysLeuSerLysIleAspProLysAlaAlaAlaIleHisPro	156
O	y	548	CATGACAAAGCCAAAGTGCGCCAGAGCTTGCACTTTTGAAGAACAGAGATCTCAT	607
			: : : : : : : : : : : : : : : : : : : : :	
D	b	157	AsnaenYrArgValIleArgAlaLeuGluIleLeuLysLeuThrGlyStnVal	176
O	y	608	AGTGAATTTCCATCCGTCACATACAGAAAGAGTGCGTCCCTTGGAGGCTCTCTG	667
			: : : : : : : : : :	
D	b	177	ThrgLu-----GlnAlaArgHisGluGluIleThrProSerProTyrr-----	190
O	y	668	AAGTCTCTCAACCCCTGTCAGCTCTTGGCTATGCTCACAGGACCTTTAGATGAGCG	727
			: : : : : : : : : : : : : : : : : : : :	
D	b	191	-----AsnLeuValMetIleGlyLeuThrMetLuarGspValIleuTyrrAspArg	207
O	y	728	TTGATTAAGAGGGTGATGATGATCATCTCTGCTGGCGCTTGGAGAGAACTAAGAGATTTT	787
			: : : : : : : : : : : : : : : : : : : : :	
D	b	208	IleAsnAlargValAspGlnMetValGluGluIleuLysIleAspGlnAlaLysLeu	227
O	y	768	CACAGACCTTAATCAAGAAATGTTTGGAAATATGCCAGACTATCAACTGTATTC	847
			: : : : : : : : : :	
D	b	228	-----TyrAspArgGlyIle	232
O	y	848	-----TTCCATCAATTTGGCTTCAGAGAAATTTCCAGAGTACCTGATCACT	892
			: : : : : : : : : : : : : : : : : : : : :	
-D	b	233	ArgAspCysGlnSerValGlnAlaIleGlyTyrIleStuMetTyrrAspTyrrLeu-----	250
O	y	893	GAGGAAAAATGCACACTGGAGACTAGTAAACCGCTTCTAAAGAA	937
			: : : : : : : : : : : : : : : :	
D	b	251	AspGlyAsnValThrLeuGluGluAlaIleAspThrLeuSary	265
RESULT 3				
MIAA_BACSU				
MIAA_BACSU	STANDARD:	PRT:	314 AA.	
O31795;				
AC	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
PT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	tRNA delta(2)-Isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPP			
DE	transferase) (Isopentenyl-diphosphate:tRNA Isopentenyltransferase)			
DE	(IPPTase) (IPPT).			
GN	MIAA.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1423;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RX	MEDLINE=98044033; PubMed=9384377;			
RA	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,			
RA	Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,			
RA	Borries R., Boutsier L., Brans A., Braun M., Birignell S.C., Bron S.,			
RA	Brouillet S., Busch C.V., Caldwell B., Capuano V.J., Carter N.M.,			
RA	Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,			

RA Dentzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrati E., Fougere D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Gallier N.,
 RA Gilm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Gutsepil G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert B., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris H., Karamata D., Kasahara Y., Klaere-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinot S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Moore D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,
 RA Preecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Kocha E., Kocha B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeder R., Scofield F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Sojido B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,
 RA Tosto V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viati A., Mamut R., Medier E., Medier H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zumbstein A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*."

RL Nature 390:249-256(1997).

CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF
 CC 2-METHYLTHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS(2)I(6)A)

CC ADACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: isopentenyl diphosphate + trna -> diphosphate +
 CC trna containing 6-isopentenyladenosine

CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).

CC
 CC DR EMBL: 299113; CAB13617.1;
 CC DR Subtilist: BG12617; miaa;
 CC DR InterPro: IPR002627; IPTP.
 CC DR Pfam: PF01715; IPTP; 1;
 CC DR ProDom: PD004674; IPTP; 1;
 CC DR TIGRfams: TIGR00174; miaa; 1;
 CC KW Transferase: Nucleotidyltransferase; trna processing; ATP-binding;
 KW Complete proteome.
 KW NP_BIND 13
 KW FT 20 ATP (POTENTIAL)
 KW SEQUENCE 314 AA: 35679 MW: 0782D3965324C1FF CRC64:

Alignment Scores:
 Pred. No.: 1,14e-23 Length: 314
 Score: 403.50 Matches: 97
 Percent Similarity: 55.56% Conservative: 63
 Best Local Similarity: 33.68% Mismatch: 29
 Query Match: 11.29% Indels: 99
 DB: 1 Gaps: 8

US-09-513-151-3 (1-2041) x MIAA_BACSU (1-314)

QY 80 CCTCTTGATGATCTCTGGGCGACCGGCGCAATTCACGCTGGCTGCAGCTA 139
 DB 7 ProvalaValleleuValgylProthralaValgylThrasnuenseerileglnleu 26

QY 140 GGCAGCGGCTGGCGGTAGATCGTACGCTGACTCATCATGAGTCTATAGAGCCTA 199
 DB 27 AlalysSerleuAsnAlaGluIleleSerGlyaspSerMetChlIleThrysglyMet 46

QY 200 GACATCATCACCAAGGTTCTGCGCCAGAGAGAGAACTGCGCGGACCACTGATGAC 259
 DB 47 AsplilegylThralalyslethrlnglInglumetgluylValProhlsleuile 66

QY 260 ACCTTGATGATCTCTGTCACCAATTCACAGTGGTGACTTCAGAAATGACCAACT 319
 DB 67 AsplileuAspPro----GlnaspeSerPheSerThrAlaAspTylGlnSerleuValarg 85

QY 320 GCTCGATGATGATATATTTCCGAGACAAATTCATGTTGTGGGAGAGCAACT 379
 DB 86 AsnlylleserGluIleAlaAsnArglyLysleuPromeTleAspIlylthrgly 105

QY 380 TATTCATGATGATCTCTGTCGTAAGAGTCTTGCAATTCACCAAGCCCGAGAGATGGGC 439
 DB 106 LeuTylrileglnSerGlnleuTyrAspTylrhrPheThr-----GluGlnAla 121

QY 440 ACTGACAAAGTATGATGACCGAAGAGCTT-----GAAAGAGAGATGCTTGA 493
 DB 122 AsnaSerProvalaPheArgGlnSerMetGlnMetAlaAlaGluArgGluLysAlaAspPhe 141

QY 494 CTTCACAAAGGCTTAAGCAGAGTGAGCCGACGAAAGGCTGCCAGCTGATCCAGATGAC 553
 DB 142 LeuHlsAlaLysleuAlaAlaAlaAspProGluAlaAlaAlaAlaIleHlsProAsn 161

QY 554 AAACGCAAGTGGCGAGAGCTTGAAGTTTGAAGAAACAGAGATCTTCATAGTGA 613
 DB 162 ThrArgArgValIleArgAlaLeuGlnIleuHlsThrSerGlyLysThrMetserGln 181

QY 614 TTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 673
 DB 182 HlsleuLysGlnGlnLysArgGln-----LeuLeu 191

QY 674 TCTAACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733
 DB 192 TyrAsnAlaValleuIleLysleuThrMetAspArgAspThrLeuTyrGlnIleuAsn 211

QY 734 AAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 793
 DB 212 GlnArgValAspLeuMetMetGlnSerGlyLeuLeuProGluVal-----Lys 227

QY 794 CCTATATACAAAGATTTTTCGGAATATACCAAGATATACACATGATGATGATGATGAT 853
 DB 228 ArgLeuTyrAspLysAsnVal-----ArgAspCysGlnSer-----IleGln 241

QY 854 TCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 913
 DB 242 AlalleglyTylrLysGlnleuTyrAlaTyrPhe-----AspGlyPheValThrleuSer 259

QY 914 ACTAGTACCACTTCTTAAGAA 937
 DB 260 AspaIaValGlnGlnleuLysGln 267

RESULT 4
 MIAA_LISMO
 ID MIAA_LISMO STANDARD: PRT: 305 AA.
 AC 08Y713;
 DT 15-JUN-2002 (Rel. 41, created)
 DT 15-JUN-2002 (Rel. 41, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE trna delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPTP
 DE transferase) (isopentenyl-diphosphate:trna isopentenyltransferase)
 DE (IPrase) (IPTP)
 GN MIAA OR LMO1294.
 OS Listeria monocytogenes.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxId=1639;
 RN (1)
 RC STRAIN=EGD-e / Serovar 1/2a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusnok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chkrikobory T.,
 RA Charbit A., Chetouani F., Couve E., Darvar A., Dehoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Fshli H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain J., Hauf J., Jackson D.,

```

0Y 77 CTACCTCTGTGATGATTTCTCGGGGGCCAGCGGCCGCAATTCACGCTGGCGGTGAC 136
      :::::::::::::::::::: ||| ::::::::::::::::::::
Db 4 IleProValIleValIleValIglProThrIalValIglYstrHsrIleuSerIleGl 23
      :::::::::::::::::::: ||| ::::::::::::::::::::
0Y 137 CTAGGCGACGGCGTCCGGCGGTGAGATGCTAGCGCTACCTACATGCAAGGTCTTGAAGC 196
      ||| :::::::::::::::::::: ||| ::::::::::::::::::::
Db 24 LeuIalYstrLysLeuAspGlyGluIleIleSerGlyAspSerMetGlnValTYrArgLy 43
      :::::::::::::::::::: ||| ::::::::::::::::::::
0Y 197 CTAAACATCACCACCAACAAAGGTTCTCGCCCAAGACAGAGAATCTCCGGCACCCATG 256
      ||| :::::::::::::::::::: ||| ::::::::::::::::::::
Db 44 LeuAspIleGlyThrIalYstrIalYstrHrProGluIleTasGluIleYstrSHAsTYrLeu 63
      :::::::::::::::::::: ||| ::::::::::::::::::::
0Y 257 ATCAGCTTGTGATTCCTGTGTGACCAATTCACAGCTGGGAGCTGCACAAATATAGCA 316
      ||| ||||| ::::| ||::| ||::|
Db 64 IleAspValThrAspPro---SerGluProPheThrIalalYstrPheGlnThrGluThr 82
      ||| ||||| ||||| ||::| ||::| ||::|
0Y 317 ACCTGCTGATTTGAAGATATTTATTTGGCCGAGACAAATATCTTATTTGTGGGAGAAC 376
      ||||| ||| ||| ||||| ||||| ||||| ||||| |||||
Db 83 ArgLysTrpIleGluThrIleHisGlnIalagLYLysLeuProIleIleValIglYlThr 1022
      ||||| ||| ||| ||||| ||||| ||||| ||||| |||||
0Y 377 AATATTACATTGATGATCTGCTGCTGGAAAGTTCTTGCTCAATACCAAGCCCGAGAGATG 436
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 103 GlyLeuYrIleGlnSerValPheTYrAspPYrAspPheGlyAsnValSerGluAspLys 1222
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
0Y 437 GGCACCTGAGAAAGTGAATTGACCGAAGATGAGACTTGAAAAAGAGATGGTCTTGACTT 496
      ::::| ||||| ||||| ||||| ||||| ||||| |||||
Db 123 Ala-----TYrArgIalagIleuGluIleuGlnIleuAsnLYstrHrLeu 136
      ::::| ||||| ||||| ||||| ||||| ||||| |||||
0Y 497 CACAAAGCCTTAGCCAGGTGAGACCCAGAAATGGCTGCCCAAGCTGCATCCATGACAA 556

```

CC use by non-profit institutions as long as its content

- 1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY. -----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to license@isb-sdb.ch).

CC EMBL: AL596168; CAC96563.1; -

DR ListList: L1N01332; -

DR InterPro: IPR002627; IPT.

DR Pfam: PF01715; IPT; 1.

DR ProDom: PD004674; IPT; 1.

DR TIGRfams: TIGR00174; miaA; 1.

KW Transferase; Nucleotidyltransferase; tRNA processing; ATP-binding;

KM Complete proteome.

FT NP_BIND 11 18 ATP (POTENTIAL).

SO SEQUENCE 305 AA; 34691 MW; 1E46250F3BA78C07 CRC64;

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
4.73e-23	395.50	51.71%	33.22%	11.06%	305	97	54	102	39	6

US-09-513-151-3 (1-2041) x MIAA_L1SIN (1-305)

77 CTACCTCTGTAGTATCTCGGGCCAGCGCCAGCAATCCACGCTGGCTGACG 136

4 IIEPVALILEVALILEVALIGYPROTHIALAVALIGLYSTHSEUSERILETHR 23

137 CTAGCCAGCGGCTCGCGGTGATGATGTCAGCGCTGATGATGATGATGATG 196

24 LENAALAYANPHEASPLGLUILEILESERGLYASPMETCLINVALTYRARGLY 43

197 CTAGACATATACCAACAGAGTTTGCCCAAGACAGACAGAAATCTGCCGCA 256

44 LEASPLILEGLYTHRALAYSLIETRRPROGLUGLMEASPOLYILEYSHSTYRLE 63

257 ATCAGCTTTGTGATCTCTTGACCAATTCACAGTGGTGGATCGAATAAGACA 316

64 ILEASVALTHRASPROALAYALPRO---PHEITHALALALYSPHEGLNALGL 82

317 ACTGCTGATGATGATGATATATTGGCCGAGACAAAATTCATTTGTTGGAG 376

83 ARGGLYLEUILEGLUSERILEHNSAENARGLYLYLEUENPROILELLEVALIG 102

377 AATTATATGATGATCTGCTGCTGGAAGTTCTTGAATRACCAAGCCCAAG 436

103 GLYLEUITYRILEGLINSEYVALPHEITYRASYRGLYPHEGLYASNALASER 122

437 GGCAGTGAAGAAGTATGACCGAAGAAGTGAAGAGATGATGCTTACTT 496

123 ALATYRATGATG---GULLEUASPOLLEUASPLYSTHRTHTLEU 136

497 CACAAGCGCTTAAGCAGGTGACCCAGCAAGAATGCTGCAAGTGCATGACAT 556

137 TTPGLMETLEUASPOLLEUASPRPOLYSSERIALAGLLEUILEHISLUNAS 156

557 CGCAAGTGGCCAGAGCTTGCAGATTTTGAAGAAGCAAGAACTCTGATGAT 616

157 ARGATGVALIILEAGLALALEUGLNUVALILEHISLEUTHRGLYSPRPHESER 176

617 CTCGACATGCAATGACGAGAGAGTGGTGGTGGAGTGGTGGAGTGGTGG 676

177 GLNVALHISHISTHLEUASNGULALATYRGLNPROLEU----- 189

677 AACCCCTTGATCTCTTGCTCATGCTGACAGGAGCTTGTAGATGAGCGCTG 736

190 -----PHELEUGLYLEUASPLLEUASPARAGLLEUENLEUITYRGLIN 206

737 AGGGTGAATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796

207 ARGVALIGLLEUENLEUENLEUENLEUENLEUENLEUENLEUENLEUEN 216

797 TATAATCAGAGAAATTTTGGAAATAGCAGAGACTATCAACATGATATC----- 847

DB 217 -----VALSERGLUASERLYSLEUTHYASPLUINHSLEUVALSPAL 231

848 -----TTCCATCAATTCAGCTTCAAGCAATTCACAGTACTGATCAGG 901

232 PROALILEARGLYLLEGLYRYSGLUENPHEUTHYRPE-----ASPLYSN 249

902 TGCACAGCTGACATGATGATGATGATGATGATGATGATGATGATGATG 937

250 SERSEUENGLUGLNUVALYSGLUENLEUENGLYLS 261

RESULT 6

MIAA_CLOAB STANDARD: PRT; 309 AA.

ID MIAA_CLOAB 097121.

AC 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPT transferase) (isopentenyl-diphosphate:tRNA isopentenyltransferase) (Iptase) (IPT).

GN MIAA OR CAC1835.

OC Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.

OX NCBI_TaxID=1488;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=21359325; PubMed=11466286;

RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q., Gibson R., Lee H.M., Dubois J., Qiu D., Hitt J., Wolf Y.I., Tatusov R.L., Sabatie F., Doucette-Stamm L., Soucaille P., Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;

RT "Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum."

RL J. Bacteriol. 183:4823-4838(2001).

CC -!- FUNCTION: Catalyzes the first step in the biosynthesis of 2-methylthio-N6-(delta(2)-isopentenyl)-adenosine (MS[2]I[6]A) adjacent to the anticodon of several tRNA species (By similarity).

CC -!- CATALYTIC ACTIVITY: isopentenyl diphosphate + tRNA = diphosphate + tRNA containing 6-isopentenyladenosine.

CC -!- SIMILARITY: BELONGS TO THE IPT TRANSFERASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to license@isb-sdb.ch).

CC -----

CC EMBL: AE007692; AAK79799.1; -

DR InterPro: IPR002627; IPT.

DR Pfam: PF01715; IPT; 1.

DR ProDom: PD004674; IPT; 1.

DR TIGRfams: TIGR00174; miaA; 1.

KW Transferase; Nucleotidyltransferase; tRNA processing; ATP-binding;

KM Complete proteome.

FT NP_BIND 9 16 ATP (POTENTIAL).

SO SEQUENCE 309 AA; 35687 MW; 9330A81B46E8AC3E CRC64;

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1e-21	378.50	54.64%	34.02%	10.59%	309	99	60	99	33	9

US-09-513-151-3 (1-2041) x MIAA_CLOAB (1-309)

83 CTGTAGTATCTCTCGGGCCAGCGCCAGCAATCCACGCTGGCTGACGCTAGGC 142

```

Db      4 11leu11le1lae1y1roth1a1a1g1y1st1h1asp1le1ser1le1y1st1le1a 23
      143 CAGCGCGTGGGCGGAGATCGTCACCGCTGACTCCATGAGCTATGAGGCGCTAGAC 202
      24 Gln1yMetAsn1yGln1u1le1ser1le1a1asp1ser1Met1le1y1st1y1Met1asp 43
      203 ATCATCCACAACAGATTCTGCGCCAGAGACAGAGATCTGCGCGCACCATGATCAGC 262
      44 11le1y1ser1a1a1y1st1h1y1s1g1u1me1t1y1s1g1y1le1y1st1h1s1h1le1u1le1asp 63
      263 TTGTGATGCTCTGTCGACCAATTCACAGCTGCTGACTGACCAATATAGACCACTGCT 322
      64 Val1a1Asp1Pro1---Ser1Gln1u1Phe1Ser1Val1Ala1Ser1Phe1y1s1y1Met1a1Gln1asn 82
      323 CTGATGAAGATATATTGGCCGAGACAAATTCCTATTGTTGGGAGAGACCAATTAT 382
      83 Ala1le1a1asp1le1r1h1er1s1er1a1g1y1st1y1P1ro1le1t1le1Val1G1y1G1y1Th1g1y1Leu 102
      383 TACATGATCTCTGCTCTGGAAGTT---CTGTCAATTCACCAAGCCCGAGAGATGGC 439
      103 Tyr1le1Asn1ser1leu1le1cys1asn1Ty1asp1he1Th1g1y1a1Ty1r1y1s1asp1Gln1a1a1Ty1r 122
      440 ACTGAAGAAGTGAATGACCGAAAAGTGAGCTTGAAGAGAGATGGTCTT---GTACTT 496
      123 Arg1Gln1Ser1Leu1-----Gln1Ala1le1a1a1y1s1p1y1s1g1y1st1y1Leu 138
      497 CACAAAGCCTTAAGCCAGGTGAGCCAGAAATGGCTGCGCAAGCTGCATCCATGACAAA 556
      139 His1Gln1Lys1Leu1y1s1asn1le1a1sp1le1asp1ser1Ty1r1y1s1y1Leu1Ty1P1ro1asn1Asp1Leu 158
      557 CGCAAGATGGCCAGAGACTTGCAGATTGTTGAAGAAACAGAGATCTCTCATGATGAAATT 616
      159 Lys1Arg1Val1le1Arg1a1le1u1Gln1Val1Ty1r1y1le1Th1g1y1st1h1le1Ser1Gln1Leu 178
      617 CTCATCGTCACATACGAGAGAGAGTGGTCCCTTGAGAGCTCTGAAATGCTCT 676
      179 -----Asn1ser1asn1Val1Asp1Leu1Ty1r1Asp1le1Pro1Ty1r1----- 189
      677 AACCCCTGATCCTTTGGCTTCATGCTGACAGGACCTTCTAGATGAGCCCTTGAGATAG 736
      190 Asn1le1His1Ty1r1Phe1le1u1Asn1Met1asp1Arg1Lys1Leu1Ty1r1Gln1a1Arg1le1Asn1Leu 209
      737 AGGTGATGATACATGCTGCTGCTGGCTCTTGAGAGAA-----CTAAGAGATTTT 787
      210 Arg1Val1Asp1le1Met1Leu1a1Arg1asn1Gly1Leu1Val1Asp1Gln1Val1le1y1s1Leu1a1Arg1asp1Met 229
      788 CACAGACGCTATATACAGAGAAATGTTTCGAAAATAGCCAGACTATCAACATGCTATC 847
      230 G1y1Ty1r1Asn1Ser1asn1Met1Gln1Ser1Met1----- 238
      848 TTCCAATCAATGCTTCAAGAGAAATTTACAGACTGATCTGATCAGGAGAAATGACACA 907
      239 ---Lys1Gly1le1Gly1Ty1r1y1s1Gln1le1Leu1Ser1Ty1r1Leu1-----Gln1Gly1Cys1le1Thr 255
      908 CTGAGACTAGTAACACGCTTCTTAAGAGAGA 940
      256 Leu1Gln1Gln1a1Val1Gln1Leu1le1y1s1Gly1 266

```

RESULT 7

MIAA_CLOPE

STANDARD;

PRT; 310 AA.

```

AC 08X1B5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPPT
DE transferase) (isopentenyl-diphosphate:RNA isopentenyltransferase)
DE (IPPTase) (IPPT).
GN MIAA OR CPE1157.
OS Clostridium perfringens.
OC Bacteria: Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.

```

```

OX NCBI_Taxid-1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-13 / Type A;
RX PubMed-11792842;
RA Shimizu T., Ohtsuka K., Hirakawa H., Oshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater."
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -1- FUNCTION: Catalyzes the first step in the biosynthesis of 2-
CC methylthio-N6-(delta(2)-isopentenyl)-adenosine (MS[2]I[6]A)]
CC adjacent to the anticodon of several tRNA species (By similarity).
CC -1- CATALYTIC ACTIVITY: Isopentenyl diphosphate + tRNA - diphosphate +
CC tRNA containing 6-isopentenyladenosine.
CC -1- SIMILARITY: BELONGS TO THE IPPT TRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
DR EMBL: AP003189; BAB80863.1; -
DR InterPro: IPR002627; IPPT.
DR Pfam: PF01715; IPPT; 1.
DR ProDom: PD004674; IPPT; 1.
DR TRFAMS: TRF000174; MIAA; 1.
KW Transferase; Nucleotidyltransferase; tRNA processing; ATP-binding;
KW Complete proteome.
FT NE_BIND 10 ATP (POTENTIAL)
FT SEQUENCE 310 AA; 35716 MW; CE421A5C406BBF CRC64;
SQ

```

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1.2e-21	310	377.50	98	66	92	37	10
Percent Similarity:		55.97%					
Best Local Similarity:		33.45%					
Query Match:		10.56%					

```

US-09-513-151-3 (1-2041) x MIAA_CLOPE (1-310)
      83 CTGTGATGATTTCTCGGGCGGACGCGGACCGCAATTCACGCTGCGCTGACATGAGC 142
      5 11leu11le1lae1y1roth1a1a1g1y1st1h1asp1le1ser1le1y1st1le1a 24
      143 CAGCGGCTGGGCGGAGATCGTCACCGCTGACTCCATGAGCTATGAGGCGCTAGAC 202
      25 Lys1y1s1Leu1asn1Gly1u1le1ser1Val1asp1ser1Met1le1y1st1y1Met1asp 44
      203 ATCATCCACAACAGATTCTGCGCCAGAGACAGAGATCTGCGCGCACCATGATCAGC 262
      45 11le1y1ser1a1a1y1st1h1y1s1g1u1me1t1y1s1g1y1le1y1st1h1s1h1le1u1le1asp 64
      263 TTGTGATGCTCTGTCGACCAATTCACAGCTGCTGACTGACCAATATAGACCACTGCT 322
      65 Phe1a1Asp1Pro1---Ser1y1s1Gln1u1Phe1Ser1Val1Ala1u1Phe1y1s1P1e1u1Th1g1u 83
      323 CTGATGAAGATATATTGGCCGAGACAAATTCCTATTGTTGGGAGAGACCAATTAT 382
      84 Lys1le1y1s1asp1le1Gln1Ser1Arg1y1Lys1Leu1ro1le1eu1Val1G1y1G1y1Th1g1y1Leu 103
      383 TACATGATCTCTGCTCTGGAAGTTCTGTGCAATTCACCAAGCCCGAGAGATGGGACT 442
      104 Tyr1le1Asn1er1-----11le1le1cys1asn1Met1asn1Phe1a1Gln1-----Ser 117
      443 GAGAAAGTGAATGACCGAAAAGTGAGCTTGAAGAG-----GAGATGGT---CTT 490
      118 Asp1y1s1asp1Gln1u1Ty1r1Arg1Gln1Leu1Gln1y1s1le1a1asn1Gln1u1H1s1G1y1Asn1Gln 137

```


Oy	491	GTACTTCAACAAAGCGCTTACGAGCTAGGTGGACCCCAAAATGGCTGCACAGCTGCATCCACAT	550
		::: :::: :::: :::	
Db	138	TyrLeuHisStuIleuMetLeuLysAspIleHisSpIeuGluSerTyrAsnSerIleHisLysPheAsn	1577
Oy	551	GACAAACGCAAAAGTGCCACGAGCGCTTGCAAGTTTTCAGAAACAGAGAACTCTCATAGT	610
		::: ::: :::: ::: :::: :::	
Db	158	AsnArgLysArgValIleArgGalaLeuGluThrTyrLysLeuThrGlyLysProPheSer	1777
Oy	611	GAATTTCTCCATCGCTCAACATACGGAAGAAAGGTGGTGTCCCTCTGGAGGTCTCTGAAAG	670
		::::	
Db	178	SerPheLysAlaLysAsnSerIle-	185
Oy	671	TTCTCTAACCGCTTCATCTCTTTGG-----CTTCATGCTGCACGAGCTTCAGAT	721
		::: :::: :::: ::: :::: :::	
Db	186	TyrGluThrProTyrAsnIleTyrTyrValLeuAsnMetAspArgAlaLysLeuTyr	205
Oy	722	GAGCGCTTGATTAAGAGGGTGGATGATCATGCTCTGCTCGGGCTCTTGAGAACTAAGA	781
		::: ::: :::: ::: :::: :::	
Db	206	AspArgIleAsnLysArgValAspIleMetPheGluLysGluLeuGluGluValLys	225
Oy	782	GATTTTCAACAGCCTTAATCAACAAGAATGTTTCGGAATAATGACGAGACTATCAACAT	841
		::: :::: :::: :::	
Db	226	AsnLeu-----LysAlaMetGlyLeuThrAspSerMetGlnSer	238
Oy	842	GGTATCTTCCAAATCAATTGGCTTCAAGAAATTTACAGCAGTACCTGATCATCGAGGAAA	901
		::: ::: ::: ::: ::: ::: :::	
Db	239	-----MetLysGlyIleGlyTyrLysGluValLeuTyrTyrLeu-----AspGlyLys	254
Oy	902	TGCACACTGGGACTAGTAAACAGACTCTTAAAGAAAGA	940
		::: :::	
Db	255	IleSerLeuGluGlnCysIleGluMetCileLysGlnIle	267

```

CC      RESULT 6
CC      MIAA_RALSO
AC      ID      MIAA_RALSO      STANDARD:      PRT;      323 AA.
CC      O8XWB0;
CC      DT      15-JUN-2002 (Rel. 41, Created)
CC      DT      15-JUN-2002 (Rel. 41, Last sequence update)
CC      DT      15-JUN-2002 (Rel. 41, Last annotation update)
CC      DE      tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPP
CC      DE      transferase) (isopentenyl1-diphosphate:tRNA isopentenyltransferase)
CC      DE      (IPrase) (IPPT) .
CC      GN      MIAA OR RSC2564 OR RS00755
CC      OS      Ralstonia solanacearum (Pseudomonas solanacearum).
CC      OC      Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
CC      OC      Ralstonia.
CC      OX      NCBI_TaxID=305;
CC      RN      [1]
CC      RP      SEQUENCE FROM N.A.
CC      RC      STRAIN=GMT1000;
CC      RX      MEDLINE=21681879; PubMed=11823852;
CC      RA      Salanoubat M., Genin S., Artiguenave F., Gonzy J., Mengenoit S.,
CC      RA      Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
CC      RA      Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
CC      RA      Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
CC      RA      Siguler P., Thebault P., Whalen M., Wincker P., Levy M.,
CC      RA      Weissenbach J., Boucher C.A.;
CC      RL      "Genome sequence of the plant pathogen Ralstonia solanacearum.";
CC      RL      Nature 415:497-502(2002).
CC      CC      -I- FUNCTION: Catalyzes the first step in the biosynthesis of 2-
CC      CC      methylthio-N6-(delta(2)-isopentenyl)-adenosine (MS[2]I[6]A1)
CC      CC      adjuvant to the anticodon of several tRNA species (By similarity) .
CC      CC      -I- CATALYTIC ACTIVITY: Isopentenyl1 diphosphate + tRNA = diphosphate +
CC      CC      tRNA containing 6-isopentenyladenosine
CC      CC      -I- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
CC      CC      -----
CC      CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      CC      the European Bioinformatics Institute. There are no restrictions on its
CC      CC      use by non-profit institutions as long as its content is in no way
CC      CC      modified and this statement is not removed. Usage by and for commercial
CC      CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      CC      or send an email to license@isb-sib.ch).

```

CC	-----
DR EMBL:	AL646070; CAD16271.1;
DR InterPro:	IIPR002827; IPTT.
DR Pfam:	Pf01715; IPTT. 1.
DR Pfam:	PD004674; IPTT. 1.
DR TIGRFAMs:	TIGR00174; maaA; 1.
KW Transferase:	Nucleotidyltransferase; tRNA processing; ATP-binding;
KW Complete proteome.	
FT NP_BIND	16 23 ATP (POTENTIAL).
SQ SEQUENCE	323 AA: 34966 MW: DP0086716A5AE1B0 CRC64;
Alignment Scores:	
Pred. No.:	1.63e-20 Length: 323
Score:	363.00 Matches: 105
Percent Similarity:	49.68% Conservative: 48
Best Local Similarity:	34.09% Mismatches: 114
Query Match:	10.15% Indels: 9
DB:	1 Gaps: 41

US-09-513-151-3 (1-2041) X MIAA-RALSO (1-323)

QY	64	CTGTGAAGGAGCCCTTACCTCTTGTAGTATCTGTGGGCGACGGGACCCGGAATTCAC	12
Db	5	ProhibitinA1ProDyA1a-ValCysLeuLeuS1ProThirA1aSerGlyS1ThirA1	24
QY	124	GCTGGCGTTGAGCTAGGCGCAGCGGCTGGGGGTAGATGCTGACCGCTCATTCATCA	18
Db	24	aAlaAlaLeuAlaLeuAlaGlnA1aGtrPrProValGluIleLeuSerMetAspSerAlaLe	44
QY	184	GCTCTATGAAGGCTTAGCAATCATTCACCCACAAAGGTTTCTGCCAAGACAGACAACTG	24
Db	44	uValTyrTrpAspMetAspIleGlyThrAlaLysProSerArgAlaGlnGlnAlaIleAl	64
QY	244	CCGGACACCAATGATACACTTGTGGATCCCTTGTGACCAATTAACAGGTGGACT	30
Db	64	aProh1n1s1eul1easPrIleIleAspProLeu---AspAlaTyrSerAlaAlaGlnPh	83
QY	304	CAGAAATGAGCAACATGCTGTGATGTAAGATATATTGGCCCGACAAAAATTCCTATGT	36
Db	83	eAlaThrAspAlaGlnAlaLeuIleGlnAlaIleAlaArgAlaIleArgLysLeuProLeuI	103
QY	364	TGTGGAGGAGCAACATTTTACATTGATGATCTGTGCTGGAATAATTCTTGCAATACCA	42
Db	103	eValGlyLeuLysMetLeuTyr-----TyrLysAlaLeuThrGlnGlyLe	118
QY	424	GCCCCAGAGATGGGCACTGGAAGAATGATGACCCGAAAGATGGAGCTTGA---AAGGA	48
Db	118	uSerAspLeuProIlyAlaAspProAlaIleAlaTylAlaGluIleAspAlaGluAlaAlaAr	138
QY	481	GGATGGCTTT---GTACTTCAAAAGCGCTTAAGCCAGGTGGAGCCCAAAATGGCTGCCAA	53
Db	138	gAspGlyTyrProAlaLeuH1sAlaTylLysLeuAlaGlnValAspProValThirAlaAlaAr	158
QY	538	GCTGCATGCSACATGACAAAGCCAAAGTGGCCAGAGCTTGCAACTTTTGAAGAAACAG	59
Db	158	gLeuH1sAlaThirAspAlaGlnArgIleGlnArgAlaLeuGluIleuTyrArgLeuThrG1	178
QY	598	AATCTGCATGAGTAATTTCCATCCATGCGCA-----CATACGGAGAGA	63
Db	178	yGlnProMetSerAlaLeuLeuAlaIleArgGluAlaGlyAlaAlaAlaPheH1sArgH1sG1	198
QY	640	AGGTGTGGTGGCCCTGGAGGCTCTGTGAAGTTCTTCAACCTTGCAATCCTTGGGCTTCA	69
Db	198	uAlaAlaAla-----AlaTyrLeuSerIleAlaLeuGlnPr	210
QY	700	TGCTGACCAAGCACTTCTAGATCAAGCGCTTGATAAAGGCTGATACATGCTGTGCTGC	75
Db	210	oAlaAspArgAlaValaLeuH1sAlaArgIleAlaGlnArgPheAspAlaMetLeuAlaG1	230
QY	760	TGGCGCTTGGAGAGACTAAAGATTTTCAACAGACGCTATACAGACAGAAATGCTTCCGA	81
Db	230	yGlyLeuLeuAspIleValGlnAlaLeuAlaGlyArgGlyGlyAspLeuSerProValLeuPr	250

QY 820 AATAGCAGACGACTATCAACATGATCTTCACATTCAGCTTCAGCAATTCACGA 879
 Db 250 oser-----1leargcysvalglytyrarglnala1trpal 262
 QY 880 GTACTGATCACTAGAGGAAATGACACTGAGACT----- 916
 Db 262 atyrleu-----aspglcyluileaspmetalatrrleuarglucinglyleala1 280
 QY 917 -AGTACACAGCTCTTAAGAAA 937
 Db 280 arthrarglnleucyslysarq 287
 RESULT 9
 MIAA_PSEPU STANDARD: PRT; 322 AA.
 ID MIAA_PSEPU STANDARD: PRT; 322 AA.
 AC O30762;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPP
 transferase) (isopentenyl-diphosphate:trna isopentenyltransferase)
 DE (IPrase) (IPPT).
 DE MIAA.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=303;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=M;
 RA Olekhovich I.N., Gussin G.N.:
 RT "Attenuation of the Pseudomonas putida type and tirpdc genes.";
 RL Submitted (Jul-1997) to the EMBL/GenBank/DDBJ databases.
 CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF
 CC 2-METHYLLHIO-N6-DELTA(2)-ISOPENTENYL-ADENOSINE (MS1211[6]A)
 CC ADACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: isopentenyl diphosphate + trna -> diphosphate +
 CC trna containing 6-isopentenyladenosine.
 CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF016312; AAB69443.1; -.
 DR InterPro: IPR002627; IPPT.
 DR Pfam: PF01715; IPPT; 1.
 DR ProDom: PD004674; IPPT; 1.
 DR TIGRFAMs: TIGR00174; miaa; 1.
 KM Transferase; Nucleotidyltransferase; trna processing; ATP-binding.
 FT NP_BIND 12 19 ATP (POTENTIAL).
 FT SEQUENCE 322 AA; 35488 MW; DIAA81AD7B32F6A CRC64;
 SQ
 Alignment Scores:
 Pred. No.: 2e-19 Length: 322
 Score: 349.00 Matches: 89
 Percent Similarity: 53.248 Conservative: 67
 Best Local Similarity: 30.388 Mismatches: 109
 Query Match: 9.768 Indels: 28
 DB: 1 Gaps: 9
 US-09-513-151-3 (1-2041) x MIAA_PSEPU (1-322)
 QY 80 CCTCTGTAGATGATTCCTCGGGCCAGGCGACGCAATTCACGCTGCGGTGACACTA 139
 Db 6 ProAlaIlePheLeuMetGlyProThrAlaIleGlyLysThrAspLeuAlaIleGluLeu 25
 QY 140 GGCACAGGCTCGGCGGTAGATCGTCACGCGCATCCATCGACGCTCTGAGGCTTA 199
 Db 140 GGCACAGGCTCGGCGGTAGATCGTCACGCGCATCCATCGACGCTCTGAGGCTTA 199

Db 26 ThrIysValLeuProCysGluLeuIleSerValAspSerAlaLeuValTyrIrrgIlyMet 45
 QY 200 GACATCATCACCACACAGGCTTCTGCCACAGACACAGAAATCTGCCGACACATGATC 259
 Db 46 AspIleGlySerAlaLysProSerLysGluIleLeuAlaAlaHisProHisArgIleLeu 65
 QY 260 AGCTTGTGGATCCTCTTGTGACCAATTCACAGCTGCTGACCTGCAATTCAGCAACT 319
 Db 66 AspIleArgAspPro--AlaGluSerTyrSerAlaAlaGlnPheArgAlaAspAlaLeu 84
 QY 320 GCTCTGATTCAGAGATATATTTGCGCGAGACAAATCTATTTGTTGGGAGAACCAAT 379
 Db 85 GluAlaMetAlaGluIleThrAlaArgGlyLysIleProLeuLeuValGlyIlyThrMet 104
 QY 380 TATTCATTCGAATCTCTGCTCTGCGAAAGTTCTTGTTCATACCAAGCCCAAGAGATGGC 439
 Db 105 LeuTyrTyrLysAlaLeuIleAspGlyLeu-----AlaAspMetPro 118
 QY 440 ACTGAGAAAGTGATGACCGGAAAGTGGAGCTGAAAGAGGAT--GGTCTGTA--- 493
 Db 119 AlaAspAlaAlaValAlaArgAlaGluLeuGluArgGlnAlaGluAlaLeuGlyLeuAlaGlu 138
 QY 494 CTTCACAAAGCGCTTAAGCCAGGTGACCCAGAAATGCTGCCAAGCTGCATCCACATGAC 553
 Db 139 LeuHisArgGlnLeuAlaGluValaAspProGluSerAlaAlaArgIleHisProAsnAsp 158
 QY 554 AAAGCAAAAGTGCCAGGAGCTTGCAGATTGTTGAAGAAACAGAAATCTCTCATAGTGA 613
 Db 159 ProGlnArgLeuIleArgAlaLeuGluValTyrArgValSerGlyGluSerMetThrAla 178
 QY 614 TTTCCTCCATCCTCAACATACGACGAGAAAGTGCGGCC-----CTGGAGGCTCTG 667
 Db 179 HisArgGlnArgIlePheAlaGluSerArgGlyAlaAspAlaGlyAlaGlyGlyHisLeu 198
 QY 668 AAGTCTCTAACCTTGCATCTTGGCTTCAAGCTGACACGACGACCTTGTAGTGAAGCC 727
 Db 199 ProTyrThrValaIleSerLeuAlaIleAlaProThrAspArgHisIleLeuHisGlnArg 218
 QY 728 TTGGATPAGAGGCTGATGATGATGATGCTGCTGCGGCTCTTGAGAGACTAGCAATTTT 787
 Db 219 IleAlaLeuArgPheSerGlnMetLeuGluGlnGlyPheValAspGluValArgSerLeu 238
 QY 788 CACAGACGCTATATATCAGAAAGATGTTTGGGAAATACGACATCATCAACTGATGATC 847
 Db 239 ArgAlaArg-----SerAspLeuHisAlaGlyLeu 248
 QY 848 -----TTCCATCAATTTGGCTTCAAGAAATTCACAGATGCTGATCATCTGAGGAAA 901
 Db 249 ProSerIleArgAlaValGlyTyrArgGlnValTyrAspTyrLeu-----AspGlyLys 266
 QY 902 TGCACACTGAGACTAGTAAACAGCTCTTAAGAAAAGA 940
 Db 267 Leu-----ThrGluAsnGluMetArgGluArgGly 276
 RESULT 10
 MIAA_THEME STANDARD: PRT; 305 AA.
 ID MIAA_THEME STANDARD: PRT; 305 AA.
 AC O9WY25;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPP
 transferase) (isopentenyl-diphosphate:trna isopentenyltransferase)
 DE (IPrase) (IPPT).
 DE MIAA OR TM0525.
 GN Thermotoga maritima.
 OS Bacteria; Thermotogae; Thermotogales;
 OC Bacteria; Thermotogae; Thermotogae (class); Thermotogales;
 OC Thermotogaceae; Thermotoga.
 OX NCBI_TaxID=2336;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB / DSM 3109;
 MEDLINE=99287316; PubMed=10360571;

RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Lither K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.,
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of *Thermotoga maritima*." ,
 RL Nature 399:323-329(1999).
 CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF 2-
 CC METHYLTHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS[2]I[6]A))
 CC ADJACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Isopentenyl diphosphate + tRNA - diphosphate +
 CC tRNA containing 6-isopentenyladenosine.
 CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AE001728; AAD35610.1; -
 DR TIGR: TM0525; -
 DR InterPro: IPR002627; IPTP.
 DR Pfam: PF01715; IPTP.1.
 DR ProDom: PD004674; IPTP.1.
 DR TIGRFAMS: TIGR00174; miaA; 1.
 KM Transferase; Nucleotidyltransferase; tRNA processing; ATP-binding;
 KM Complete proteome.
 FT NP_BIND 8 15 ATP (POTENTIAL).
 SO SEQUENCE 305 AA; 35926 MW; 59C3C845AEC64AC3 CRC64;
 Alignment Scores:
 Pred. No.: 2,366-19 Length: 305
 Score: 348.00 Matches: 86
 Percent Similarity: 53.66% Conservative: 68
 Best Local Similarity: 29.97% Mismatches: 105
 Query Match: 9.73% Indels: 28
 DB: 1 Gaps: 8
 US-09-513-151-3 (1-2041) x MIAA_THEMEA (1-305)
 QY 77 CTACCTCTGTGATGATTCGCGGCCACGGCCGGAATCCACGCTGCGCTTGACG 136
 DB 1 MetLysIleAlaIleValGlyProThrAlaValAlGlyLysThrAspIleMetIleGlu 20
 QY 137 CTAGGCCAGCGCGCTGCGGGTGAATCGTCAGCTCATGCAAGGTATGAGGC 196
 DB 21 ValGysGluGluIleGlyAlaGluIleIleSerMetAspSerArgGlnIleTyrGly 40
 QY 137 CTAGCATATCATCCACACAGAGTTTCCGCCACAGACAGAAATGCGCGCACCATG 256
 DB 41 MetAspIleGlyThrAlaLysProThrProGluGlnArgLysArgValIleHisMet 60
 QY 257 ATCACTTGGTGATGCTCTTGACCAATTCACAGTGGTGCATTCGAAATAGACGA 316
 DB 61 IleAspIleIleAspPro---AspGluTyrTyrAsnAlaPheMetTyrArgLysAspSer 79
 QY 317 ACTGCTGTGATGAAGATATATTCGCCGACACAATTCATGTTGTGGAGAGAAC 376
 DB 80 LeuArgAlaMetGluAspValLeuArgArgLysIleProValTyrValGlyGlyThr 99
 QY 377 AATTATTTCATGTAATCTCTGCTCGGAAGTCTTGTCAATACCAACGCCAGAGATG 436
 DB 100 GlyLeuTyrAlaAspAlaLeuVal---ArgGlyIlePheGluGlyValProAlaAspGlu 118
 QY 437 GGCACGTAGAAAGATGATGACCGAAATGACGCTGAAAAGAGAGATGCTTACTT 496
 DB 119 AsnIleArgLysGluLeuArg-----GluLeuGlnArgArgGluProGlyIleLeu 135

QY 497 CACAAAGCGCTTAAGCCAGCTGGACCCAGAAATGCGTCCAGACTGCATCCACATGACAA 556
 DB 136 ArgLysMetLeuGluIleuAspProGluAlaAlaThrArgIleHisProAsnAspLeu 155
 QY 557 CCCAAAGTGGCCAGAGCTTGCAGAGTTTGTGAAGAAGACGAAATCTCATGATGTAAT 616
 DB 156 LysArgThrIleArgAlaLeuGluValTyrMetLysThrGlyArgArgIleSerGluLeu 175
 QY 617 CTCATCGTCAACATACGGAAGAGGTGGTCCCTTGAGAGTCTCTGATCTCT 676
 DB 176 -----GlnLysGluAlaLysGlyAspAsp-----ArgPhe----- 185
 QY 677 AACCCCTGATCCTTGGCTTCATGCTGACACGCGCTCTAGATGAGCGCTTGATAG 736
 DB 186 -----PheIleIleValIleuThrArgGlyGlyLeuTyrGluArgIleAsnLys 203
 QY 737 AGGGTGATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
 DB 204 ArgValAspLysMetIleGluMetGlyLeuValAspGluValLysArgLeuGlyMet 223
 QY 797 TATTAATCAGACAAATGTTTGGAAAAATACCGACGACTATCAACATGATTCATCA 856
 DB 224 GlyTyrSerLysAspLeuAsnSer-----MetLysThr 234
 QY 857 ATGGCTTCAAGCAATTCACAGATACCTGATCAGTACAGGAAATGCACACTGAGACT 916
 DB 235 IleGlyTyrLysGluValIleAspTyrLeu-----GluGlyLysTyrAspPheAspLys 252
 QY 917 AGTAACAGCTCTTAAGAA 937
 DB 253 MetValHisLeuIleLysArg 259
 RESULT 11
 MIAA_CHLTR STANDARD: PRT; 314 AA.
 AC 084771;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPT
 DE transferase) (Isopentenyl-diphosphate:tRNA isopentenyltransferase)
 DE (Iprase) (IPTP).
 GN MIAA OR CT766.
 OS Chlamydia trachomatis
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D/UM-3/Cx;
 RA MEDLINE=99000809; PubMed=9784136;
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
 RA Mitchell W.P., Olinger L., Yatsunov R.L., Zhao Q., Koonin E.V.,
 RA Davis R.W.;
 RT "Genome sequence of an obligate intracellular pathogen of humans:
 RT Chlamydia trachomatis." ,
 RL Science 282:754-759(1998).
 CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF
 CC 2-METHYLTHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS[2]I[6]A))
 CC ADJACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Isopentenyl diphosphate + tRNA - diphosphate +
 CC tRNA containing 6-isopentenyladenosine.
 CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AE001349; AAC68361.1; ALT_INIT.
 DR InterPro: IPR002627; IPTP.

QY 614 TTTCATCATGTCACATAGAGAGGAGGTCCTCCCTGGAGGTCCTCTGAAGTTC 673
 Db 178 Leu-----ThrlGlnLysGlyGluAlaLeu-----ProTyrAspPhe 190
 QY 674 TCTAACCTTCGATCCTTGGCTTCATGCTGACACGAGCATTCATAGAGCGCTTGAT 733
 Db 191 ValGlnPheAlaIle-----AlaProGlnAspArgHisValLeuHisGluArgIleGlu 208
 QY 734 AAGAGGCTGATGACATGCTGCTGCTGCTGCTGGAGAGACTAAGATTTTACAGA 793
 Db 209 GlnArgPheHisLysMetIleGluLeuGlnAlaGluValGlnLysLeuTyrAla 228
 QY 794 CGCTATATACAGAGAGATTTTCGAGAAATACGACAGACTATCAACATGATCTCCAA 853
 Db 229 Arg-----GlyAspLeuAsnIleAsnLeuProSer-----IleArg 240
 QY 854 TCATATGCTTCAGAGAAATTTTCACGAGTACTG----- 886
 Db 241 CysValGlyTyrArgGlnMetTrpGluTyrLeuGlnGlyAspTyrAlaTyrGluGluMet 260
 QY 887 ATCAGTACGAGGAAATGCACACTGAGACTAGTACACGCTTCTAAAGAA 937
 Db 261 IlePheArgGlyIleCys-----AlaThrArgGlnLeuAlaLysArg 274
 RESULT 15
 MIAA_VIBCH STANDARD: PRT: 315 AA.
 ID MIAA_VIBCH 09KV12;
 AC 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPT
 transferase) (isopentenyl-diphosphate:tRNA isopentenyltransferase).
 DE (Ippase) (IPPT).
 GN MIAA OR VC0346.
 OS Vibrio cholerae.
 OX Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 NCBI_taxonomy:666;
 RN NCBI_taxonomy:666;
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 Dodson R.V., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
 Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 Ermolaeva M.D., Vamathevan J., Baas S., Qin H., Dragoi I., Sellers P.,
 McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 Fraser C.M.;
 RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 cholerae".
 RT Nature 406:477-483(2000).
 RL Nature 406:477-483(2000).
 CC -I- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF 2-
 METHYLTHTIO-NE-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS1211(6IA))
 CC ADJACENT TO THE ANTICODON OF SEVERAL tRNA SPECIES (BY SIMILARITY).
 CC -I- CATALYTIC ACTIVITY: isopentenyl diphosphate + tRNA = diphosphate +
 CC tRNA containing 6-isopentenyladenosine.
 CC -I- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL: AE004123; AAF93519.1; -
 DR TIGR: VC0346; -
 DR InterPro: IPR002627; IPPT.
 DR Pfam: PF01715; IPPT: 1.
 DR ProDom: PD004674; IPPT: 1.

DR TIGRFA: TIGR00174; mia: 1.
 KW Transferase; Nucleotidyltransferase; tRNA processing; ATP-binding;
 KM Complete proteome.
 FT NP_BIND 13 20 ATP (POTENTIAL).
 SQ SEQUENCE 315 AA; 35515 MW; 0114D280868CD876 CRC64;
 Alignment Scores:
 Pred. No.: 2, 93e-18 Length: 315
 Score: 334.00 Matches: 99
 Percent Similarity: 49.86% Conserved: 76
 Best Local Similarity: 28.21% Mismatches: 116
 Query Match: 9.34% Indels: 60
 DB: 1 Gaps: 14
 US-09-513-151-3 (1-2041) x MIAA_VIBCH (1-315)
 QY 65 CTGCAAGAGACCTCCTTCTGATG---ATTTCGGGCGCAGCGGCAATGCC 121
 Db 1 MetThrGlnLysLeuProLeuAlaLeuPheLeuMetGlyProThrAlaSerGlyThr 20
 QY 122 AGCGTGGCTGCGAGTACGCGCGCGCGCGGCGGATGCTGACCGCTGATCCATG 181
 Db 21 AspLeuAlaIleArgLeuArgGlnLysTyrProValGlnIleIleSerValAspSerAla 40
 QY 182 CAGGTCTATGAGGCGCTACACATCATCACCACCAAGAGTTTCCGCAAGAGCAATC 241
 Db 41 LeuIleTyrArgGlyMetAspIleGlyThrAlaLysProAspAlaGlnGluLeuAlaLeu 60
 QY 242 TCCGCGCACCACATGATCAGCTTGTGATTCCTTGTGACCAATATACAGTGGTGCAC 301
 Db 61 AlaProHisArgLeuIleAspIleLeuAspProSerGluAla---TyrSerAlaAlaAsp 79
 QY 302 TTCAGAAATAGACCAACTGCTGATGAGATATATTTGCCGAGACAAATTCCTATT 361
 Db 80 PheArgArgAspAlaLeuLysGluMetAlaAspIleValAlaGlnGlyLysIleProLeu 99
 QY 362 GTTGTGGAGGAGAACCAATTATTTACATTGATCTCTCTCTGCAAACTTTTGCAATACC 421
 Db 100 LeuValGlyGlyThrMetLeuTyrPheLysAlaLeu-----LeuGlnGlyLeu 115
 QY 422 AAGCCCCAGAGATGGGCACTGAGAAATGATGACGGAAGTGGAGCTTGAAGAGAG 481
 Db 116 SerProLeuPro---AlaAlaAspProValIleArgGlnGlnIleGluGlnAlaGlu 134
 QY 482 -----GAGGTCTCTGATCTTACATTCACAAACGCTAAGCAGGTGACCCAGAAATGCTGCC 535
 Db 135 LysLeuGlyTyrGlnAlaLeuHisAspGlnLeuGlnGlnIleAspProValSerAlaGln 154
 QY 536 AAGCTGATCCACATGACAAAGCAAGAGTGGCAGAGCTTGCAAGTTTGTGAAGAAACA 595
 Db 155 ArgIleHisProAsnAspProGlnArgLeuSerArgAlaLeuGlnValTyrArgIleSer 174
 QY 596 GGAATCTCATAGTAAATTTCTCCATCGTCAACATACGAGAAAGTGGTGGTCCCT 655
 Db 175 GlyLysThrLeuThrGluLeu-----ThrGlnThrLysGlyAlaIle 189
 QY 656 GAGGTCTCTGAGTCTCTACCCCTTGATCTTGGCTTCATGCTGACGAGGAGT 715
 Db 190 -----ProTyrArgValLeuGlnPheAlaIle-----AlaProLysGlnArgAlaGln 205
 QY 716 CTAGATGAGCGTTGGATAGAGGGTGGATGACATGCTGCTGGCTTTGGAGGAA 775
 Db 206 LeuHisArgArgIleGlnLeuArgPheGlnLysMetValGlnSerGlyPheGlnGluGln 225
 QY 776 CTAAAGATTTTCAACAGACGCTATATACGAGAAATGTTTCGAAATATGACGAGACTAT 835
 Db 226 ValLysAlaLeuTyrAlaArg-----AspAsp 234
 QY 836 CAACATGATATCTTCATCAAT---GGCTTCAGAGAAATTTCCAGAGTACTG 886
 Db 235 LeuHisProAspLeuProSerIleArgCysValGlyTyrArgGlnMetTrpTyrLeu 254
 QY 887 ATCACTAGAGGAAATGCACACTGGAG-----ACTAGTAAAC 922

```
Db 255 -----AspGlyHisGlyThrLeuAspGluAlaIleTyrArgGlyIleCysAlaThrArg 272
QY 923 CAGCTTCTAAGAAAGACCTGCTCCATTGTCCTAGGCTTAGAGTACT 982
Db 273 GluLeuAlaLysArg-----GlnIleThr 280
QY 983 GATGCTCTGAAAGTGGAGAGACTCTGT-----CTTGAACCTGCTCTT 1024
Db 281 TrrLeuArgSerTrrAspAspLeuThrTrrLeuAspSerGluAsnValAspGluAlaVal 300
QY 1025 GAATCGTCGCAAGTTTCATCCAGGCCACAG 1057
Db 301 GluThrLeuSerAsnAlaIleAlaSerAsnGlu 311
```

Search completed: April 21, 2003, 18:49:17
Job time : 32.423 secs